Experiments on intertemporal choices and belief change

Sun, Chen

Experiments on Intertemporal Choices and Belief Change

A dissertation presented

by

CHEN SUN

to

The CentER Graduate School

Tilburg University

Tilburg, The Netherlands

July 2018

Experiments on Intertemporal Choices and Belief Change

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 19 september 2018 om 16.00 uur door

Chen Sun

geboren te Shanghai, China

PROMOTIECOMMISSIE:

PROMOTORES:	Prof. dr. J.J.M. Potters
	Prof. dr. A.H.O. van Soest
	Prof. dr. G. van de Kuilen
OVERIGE LEDEN:	Prof. dr. M. Abdellaoui
	Dr. E. Cettolin
	Prof. dr. K.I.M. Rohde
	Prof. dr. S.T. Trautmann

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisors, Jan Potters, Arthur van Soest and Gijs van de Kuilen, without whose guidance this dissertation would not be possible. Jan is an excellent supervisor who is inspiring and patient. With his encouragement, I pursued my research interests in intertemporal choices and behavior change. In our hour-long meetings or occasional talks, Jan was always seriously answering all my questions. I received a great number of helpful comments from him on my research plan, experimental designs and presentations. His broad vision and good economic intuition makes him a role model for me. Arthur provided me much guidance on econometric methods. I am deeply impressed by his preciseness and knowledgeability which had a great impact on my research attitude. Gijs provided me useful advice and kind support. Without him I would not have had so much great opportunities to know people who have the same research interests as me. I was so lucky to have all of them illuminating my road to becoming a qualified researcher.

I would like to express my sincere gratitude to the committee members: Mohammed Abdellaoui, Elena Cettolin, Kirsten Rohde and Stefan Trautmann for their careful reading and constructive feedback which significantly improved the dissertation.

It was my great pleasure to have insightful discussions with some colleagues about my research. Wieland Müller shared his critical and helpful comments on two chapters of my dissertation. Xu Lang and I had numerous brainstorms during coffee breaks and on our way home. I appreciate their help a lot.

Many other researchers have contributed their knowledge and expertise to the improvement of my dissertation. Those include but are not limited to: Aurelien Baillon, Jan Boone, Stephen Cheung, Patricio Dalton, Eric van Damme, Mark Dean, Sebastian Ebert, Eleonora Freddi, Reyer Gerlagh, Georg Granic, Glenn

Harrison, Boris van Leeuwen, Chen Li, Guy Mayraz, Bin Miao, Peter Moffat, Jens Prüfer, David Schindler, Karl Schlag, Sebastian Schweighofer-Kodritsch, Yuehui Wang, Bert Willems, Xue Xu, Yilong Xu, Songfa Zhong and Xiaoyu Zhou. I also benefited from discussing econometrics with Yifan Yu and Bo Zhou.

I am grateful to my colleagues and friends who made my life at Tilburg a lot of fun: Cansu Aslan, Sebastian Dengler, Di Gong, Victor Gonzalez Jimenez, Chen He, Bas van Heiningen, Dorothee Hillrichs, Yufeng Huang, Hong Li, Hao Liang, Manxi Luo, Mingye Ma, Renata Rabovič, Julius Rüschenpöhler, Gyula Seres, Lei Shu, Ruixin Wang, Xiaoyu Wang, Jianxing Wei, Xingang Wen, Huaxiang Yin, Zhiyu Yu, Baiquan Zhai, Xinyu Zhang, Yi Zhang, Kun Zheng, and many others.

My experiences of attending academic conferences contributed to the dissertation as well. I am very grateful to the support from the departmental and CentER staff, especially Korine Bor and Cecile de Bruin. With their help, going for conferences became very easy. I also thank the Netherlands Organisation for Scientific Research (NWO) for their financial support to my work.

Finally, I would like to express my heartfelt gratitude to my parents for their continuous encouragement and unconditional love during my seven years at Tilburg. I also thank my girlfriend Yuxin Yao for her companionship (㚘) and being interested in and supportive to my research.

TABLE OF CONTENTS

Page

List of Tables

LIST OF FIGURES

1 INTRODUCTION

This dissertation comprises three essays in economics. They concern two topics: intertemporal choices and belief change. All of the three essays make use of experimental methods.

Many economic decisions involve outcomes at different points in time. Kids decide how to finish a jar of sweets over a month. Senior students decide whether to study in a graduate program or to work. Employees decide how much to invest in pension accounts. Typical questions faced by those decision-makers are whether they want a smaller sooner reward or a larger later reward, and whether they want a big reward at a time or small rewards over a period of time.

Economic experiments on intertemporal choices often start from choices between monetary rewards. A prediction of the standard consumption-savings model is that people should always discount small (compared to lifelong wealth) monetary rewards at the market interest rate. However, in almost all experiments the majority of subjects do not behave in that way. Therefore, it is still interesting to know how people make intertemporal choices over monetary rewards.

In Chapter 2, my coauthor and I investigate one aspect of the topic: how choices change with the size of outcomes, usually called the magnitude effect. The magnitude effect is an established pattern of intertemporal choices, dating back to Thaler (1981), but most studies only looked at choices between single dated rewards, such as choosing between receiving €20 tomorrow and receiving €30 in five months. As a result, we only knew the magnitude effect as a whole. But most economic decisions involve a more general case: choices between intertemporal profiles, such as choosing between receiving €20 tomorrow and €20 in five months and receiving €28 tomorrow and €10 in five months. To predict how such choices change with the size of outcomes, we need knowledge about channels of the magnitude effect: when outcomes are larger, whether people are more patient, and whether rewards at different points in time are more fungible.

In order to examine channels of the magnitude effect, we use the Convex Time Budget method introduced by Andreoni and Sprenger (2012). The method requires subjects to allocate a total budget into two dates, given an interest rate for money allocated to the later date. The method allows us to estimate the discount rate and the utility curvature simultaneously, and hence we are able to disentangle the magnitude effect into two channels.

We find a significant magnitude effect in intertemporal allocation tasks: the budget share allocated to the later date increases with the size of the budget. This effect does not depend on whether the sooner reward is paid in the present or in the future, implying that the factors which drive the present bias cannot fully account for the magnitude effect. At the aggregate level as well as at the individual level, we find magnitude effects both on the discount rate and on intertemporal substitutability (i.e. utility curvature). The latter effect is consistent with theories in which the degree of asset integration is increasing in the stake.

Chapter 3 is motivated by a question raised during the writing of Chapter 2. The CTB method used in Chapter 2 requires parametric assumptions of the utility function being measured, and hence our decomposition of the magnitude effect also relies on parametric assumptions. Though we have checked a few popular specifications showing the robustness of our results, I was always asking a question: is it possible to study the magnitude effect (and other aspects of intertemporal choices) in a completely parameter-free way?

In Chapter 3, I develop a simple method for measuring intertemporal preferences: directly measuring preferences over intertemporal profiles. The method is parameter-free and independent of time horizon effects, and hence researchers do not need to assume the form of the utility functions or that of the discount function. It requires weak assumptions on preferences to be measured, and hence can be used to test a wide range of models.

By applying the method, I test eight models of intertemporal choices. Those models provide different predictions on two properties of preferences: how choices change with the stake (the magnitude effect), and whether people are overly impatient when it is possible to have all money on the earlies date (the all-sooner effect).

Regarding the magnitude effect, I find that utility curvature is smaller for higher stakes, but no evidence shows that (generalized) discount factors change with the stake. Regarding the all-sooner effect, I do find evidence that people are overly impatient when a pure sooner reward is available. I then propose a simple model which captures these two effects to facilitate parametric estimation in future studies.

The other topic of this dissertation is belief change. In economic theories, when uncertainty is present, choices are based on beliefs, but do choices in turn have an effect on subsequent belief formation? For instance, John chooses yoghurt over ice cream because he believes that the health benefits of yoghurt over ice cream overweighs its disadvantage in deliciousness. Does this choice per se (rather than new information obtained from eating the yoghurt or from reading health magazines) make John more believe in a large health benefit of yoghurt?

I address this question in Chapter 4. I perform an experiment to study choice-induced belief change in an individual decision-making context. After being presented with noisy signals about the values of two options, subjects are randomly assigned to one of the three treatments: making a choice between the two options, receiving a random option, or possessing no option. Then they are presented with more signals and are asked to estimate the values of the two options.

I find no significant differences between the distributions of estimates in the three groups, irrespective of whether belief elicitation is incentivized or not, and irrespective of whether previous signals are perfectly accessible or not. This suggests that making a choice does not lead to an economically meaningful effect on subsequent belief formation in my setting.

2 MAGNITUDE EFFECT IN INTERTEMPORAL ALLOCATION TASKS

The prediction of the standard consumption-saving model, that people always discount an income at the market interest rate, has been found to be inconsistent with empirical results.[1] One important anomaly, dating back to Thaler (1981), is the magnitude effect: people appear less patient when choosing among smaller rewards than when choosing among larger rewards. A deeper understanding of this anomaly will help to lay a more solid foundation for the research of intertemporal choice and related applications.

In this paper, we investigate whether the magnitude effect on time preferences can be observed in intertemporal allocation tasks, and if so, whether the magnitudes impact intertemporal preferences through the present bias, the discount rate or the atemporal utility function.

Different channels of the magnitude effect have different implications for choices in intertemporal allocation tasks. If the discount rate is smaller for larger outcomes, when the total budget increases, people will appear more patient in all choices. If the utility curvature is smaller for larger outcomes, when the total budget increases, choices will be more sensitive to interest rates. If the present bias is smaller for larger outcomes, when the total budget increases, choices made between today and a future date will be less different from choices made between two future dates.

Several experiments on time preferences have reported a magnitude effect.[2] Though most early studies are based on hypothetical decisions, there are also some real-stake experiments that found a magnitude effect (Holcomb and Nelson,

---

[1] To be more precise, people discount at the market rate unless the borrowing constraint is binding.

[2] Frederick et al. (2002, Section 4.2.2) summarized the early literature on the magnitude effect of time preferences. Andersen et al. (2011) also reviewed the more recent literature.

1992; Kirby, 1997; Kirby, Petry and Bickel, 1999; Andersen, Harrison, Lau, and Rutström, 2013; Halevy, 2015). In this literature, little efforts are made to explore the channels of the magnitude effect. This is mainly because most studies employed a *single-reward task*, in which a subject can only get one reward, either on a sooner date or on a later date. With a single-reward task, one cannot disentangle different channels and can only attribute all effects to one aggregate measure, the monetary discount rate.

We are interested in the following question: Is the magnitude effect mainly driven by the factors which drive the present bias, does the magnitude affect choices through the long-run discount rate, or does it affect choices through intertemporal substitutability (utility curvature)? To disentangle these channels is interesting for at least two reasons. First, the knowledge about how the stake affects intertemporal choices in different ways is important for establishing deeper and better-founded descriptive theories of intertemporal decision making. Second, omitting a channel of the magnitude effect in an empirical study or in policy making may lead to misspecified models and biased estimates and predictions.

We employ the Convex Time Budget (CTB) method introduced by Andreoni and Sprenger (2012). It allows subjects to form a portfolio of a sooner reward and a later reward given a budget constraint. The possibility for subjects to make interior choices (and not only corner choices as in most binary choice tasks) enables the researcher to simultaneously identify the discount rate and the intertemporal substitutability.[3]

[3] Abdellaoui et al. (2013a) provided another method for measuring intertemporal preferences parametrically. Their method identifies utility curvature from marginal utilities for different quantities on the same date, while the CTB method identifies utility curvature from sensitivities of choices to interest rates. They are equivalent if the true model is with a stationary period utility function and a magnitude-independent discount function, as assumed in our paper. If the condition is not satisfied, the former method is better at measuring curvature of a period utility function, and the CTB method is better at measuring sensitivity of choices to interest rates (or elasticity of intertemporal substitution). A discussion about the rationality of subjects in the CTB task is provided in Appendix 2.A.

The design of our experiment has three main features. First, all subjects receive equal amounts of participation fees on the sooner date and on the later date regardless of their choices, and the payment conditions are constant across time. Thus, the transaction costs and the trustworthiness of the payments are equalized across periods, and these confounding factors are controlled for. Second, we implement two treatments. In one treatment subjects allocate between today and four weeks later, while in the other treatment subjects allocate between four weeks later and eight weeks later. This allows us to assess whether the magnitude effect is driven by the same factors that drive the present bias. Finally, by assuming a simple yet popular model, the CTB method allows us to identify the discount rate and the atemporal utility function simultaneously. As a result, we are able to disentangle the channels of the magnitude effect.

We find evidence of the magnitude effect, irrespective of whether or not a front-end delay is present, suggesting that the factor which drives the present bias cannot fully explain the magnitude effect. The magnitude effect is decreasing in the magnitude. At the aggregate level as well as at the individual level, we find magnitude effects both on the discount rate and on intertemporal substitutability. Both channels have considerable impacts on predicted choices. We find that the latter effect is not the same as the magnitude effect on risk attitudes found in previous studies, and hence it might be problematic to correct for the curvature of utility functions by risk attitudes.

The remaining part of this paper is structured as follows: We introduce our experimental design in Section 2.1. In Section 2.2 we formulate our hypotheses. In Section 2.3, we investigate non-parametrically the magnitude effect and its relation with the present bias. We explore the channels by parametric estimation both at the aggregate level and at the individual level in Section 2.4. In Section 2.5, we discuss the interpretations of our findings. We draw conclusions in Section 2.6.

2.1 Experimental Design

2.1.1 *The Convex Time Budget Method, Parameters and Implementation*

The foundation of our experimental design is the Convex Time Budget method introduced by Andreoni and Sprenger (2012). The method consists of a set of intertemporal allocation tasks: in each *decision* subjects are asked to allocate N tokens to two dates, t days from today and $(t+\tau)$ days from today. Each token allocated to t is worth P_t euro, while each token allocated to $(t+\tau)$ is worth $P_{t+\tau}$ euro. Suppose a subject allocates n_t tokens to the sooner date and $n_{t+\tau}$ to the later date, the amount of the sooner reward will be $z_t = P_t \cdot n_t$ euro and the amount of the later reward will be $z_{t+\tau} = P_{t+\tau} \cdot n_{t+\tau}$ euro.

Choices are subject to the budget constraint, $n_t + n_{t+\tau} \leq N$, and the non-negativity constraints, $0 \leq n_t, n_{t+\tau} \leq N$. They are told that they can allocate any number of tokens they like to one of the two dates. Examples of both corner choices and interior choices are given to remove any hesitation in making either type of choices.

Decisions with the same total budget, N, are grouped in one *decision form*, which is displayed on one page. There are seven decisions in one decision form. The return to each token allocated to the later date is fixed as $P_{t+\tau} =$ €0.20, while the return to each token allocated to the sooner date is varied and takes the values $P_t =$ €0.20, €0.19, €0.18, €0.17, €0.16, €0.15, and €0.14. Hence, those decisions imply seven gross interest rates, R =1, 1.05, 1.11, 1.18, 1.25, 1.33, and 1.43, respectively, over a period of τ days. The constraints can be rewritten as

$$R \cdot z_t + z_{t+\tau} \leq m$$

$$z_t, z_{t+\tau} \geq 0$$

where m is the total budget and $m = P_{t+\tau} \cdot N$.

We implement the CTB method by a zTree program (Fischbacher 2007). Figure 2.1 shows the interface of a typical decision form. Each decision takes a row. Decisions can be made by scrolling the bars. Once an adjustment is made for one decision, the amounts of the sooner reward and of the later reward in that decision are automatically calculated and displayed.

To avoid any possible effects of initial values, the amounts of rewards are initially blank. Decisions cannot be submitted until all the scrollbars have been adjusted at least once.

2.1.2 Procedures

There are two parts in our experiment. Part I consists of five decision forms, with N =100, 200, 300, 400, and 800. The order is randomly drawn for each subject. Subjects can move to a specific decision form by clicking the button with the corresponding number. One can go to any decision form at any time, regardless of whether the current decision form is completed. Decisions are automatically stored when one switches to another decision form. This makes comparisons across magnitudes very easy to the subjects in case they would want to make such comparisons. Decisions can only be submitted when all the 35 decisions in the five decision forms are completed.

We randomly assign subjects to one of two treatment groups. In the Present Group, the sooner date is today while the later date is four weeks from today, i.e. $t = 0$ and $\tau = 28$. In the Delayed Group, the sooner date is four weeks from today while the later date is eight weeks from today, i.e. $t = 28$ and $\tau = 28$. Comparing the two groups enables us to check if there exists a present bias on average, and more importantly, if there exists a magnitude effect when no rewards are available in the present.

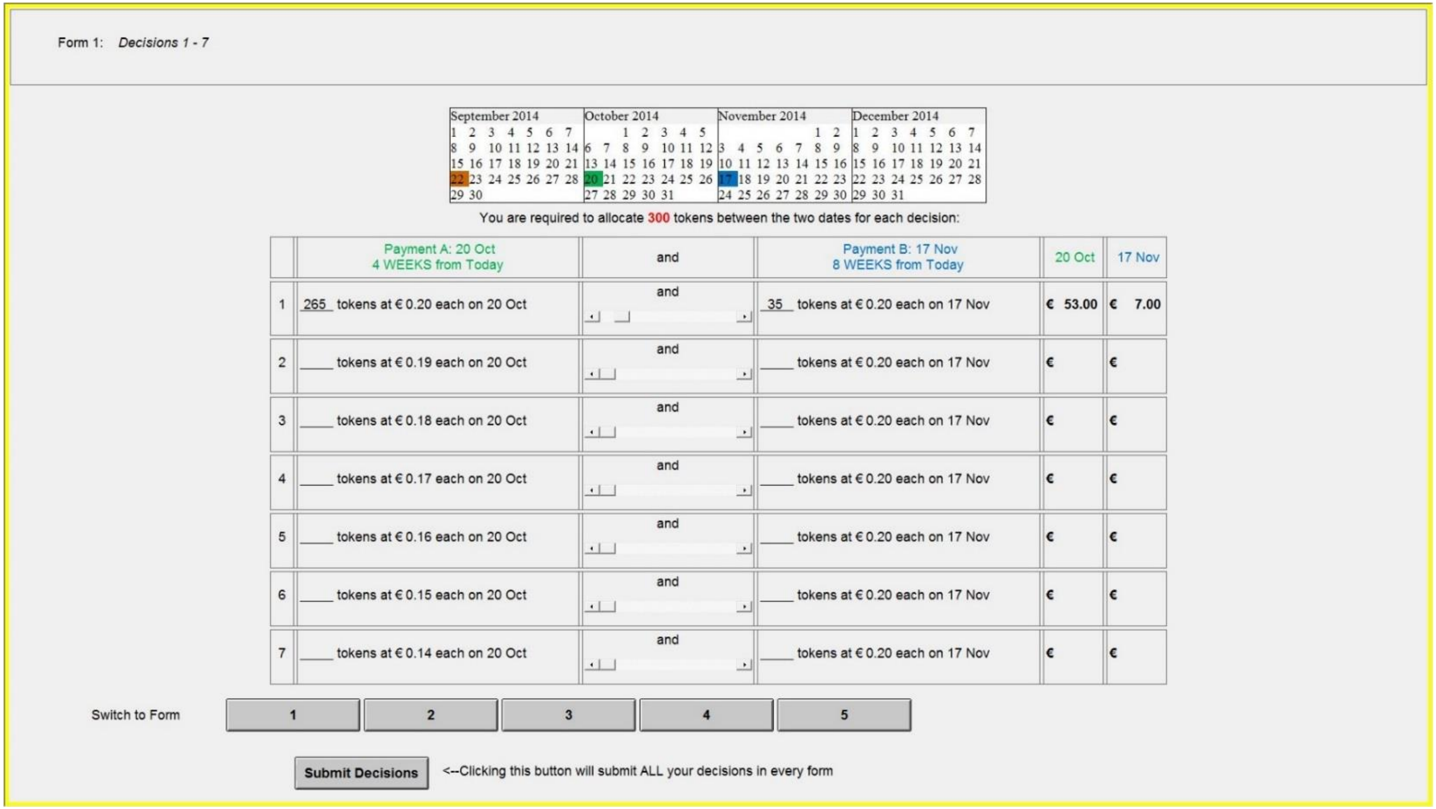


Figure 2.1: Interface of a Typical Decision Form in Part I

Part II is composed of an extended CTB decision form with seven decisions. Subjects are asked to allocate 400 tokens to three dates, today, four weeks from today and eight weeks from today. One additional restriction is imposed, depending on which group one is in. A subject in the Present Group can allocate either 0 or 200 tokens to eight weeks from today; she cannot choose other numbers. But she is still free to allocate any number of tokens between today and four weeks from today. Similarly, a subject in the Delayed Group can allocate either 200 or 400 tokens to today. She is still free to allocate any number of tokens (if there remains some) between four weeks from today and eight weeks from today. The restrictions and the returns to one token allocated are shown in Table 2.1.

Table 2.1: Restrictions on the Number of Tokens and Returns to One Token Allocated to a Specific Date in Part II

Group		Today	Four weeks from today	Eight weeks from today
Present	Returns to one token	€0.20, €0.19, €0.18, €0.17, €0.16, €0.15, €0.14	€0.20	€0.26
	Restriction on the number of tokens	No restriction	No restriction	0 or 200
Delayed	Returns to one token	€0.08	€0.20, €0.19, €0.18, €0.17, €0.16, €0.15, €0.14	€0.20
	Restriction on the number of tokens	200 or 400	No restriction	No restriction

The additional date (eight weeks from today for the Present Group or today for the Delayed Group) is accompanied with a very high return for the Present Group and a very low return for the Delayed Group, so that subjects are induced to allocate 200 tokens to this additional date.[4] If they do so, the remaining task is equivalent to the one with a total budget of 200 tokens in Part I. This characteristic makes the two decision forms comparable.

The purpose of Part II is to test the time separability of intertemporal preferences. One alternative hypothesis is that a subject in the Delayed Group may allocate less to the sooner date if she has allocated a large amount of money to an even sooner date, since the desire for extra consumption has already been partly satisfied. A similar hypothesis applies to the Present Group: a subject in the Present Group may allocate less to the later date if she has already allocated a large amount of money to an even later date, since the guilt for not saving has been partly released. If preferences are time non-separable, the use of a model with a time separable preference is more likely to be problematic. Thus, we want to test the hypothesis of time separability before we perform parametric estimation with a time-separable model.

We do not directly give a fixed reward on the additional date. This is because a fixed reward might be mentally isolated from the allocation task due to narrow bracketing, and hence the test of time separability in the allocation task may be invalid.

At the end of the experiment, subjects were asked to finish a questionnaire. As in previous studies with the CTB method, we asked about subjects' typical expenditures in one week. The average response was €55.22 per week or €7.89 per day.

2.1.3 Experimental Payments

The payments are composed of two parts. First, all subjects receive a €5 *participation fee* on each of the two dates scheduled in Part I. Second, each subject has a 10% chance to receive *earnings from decisions*. Before the

[4] In fact, only nine out of 203 subjects selected a different number than 200 to the additional date, which involved 41 (2.9%) out of 1415 decisions.

experiment starts, each subject is randomly given a lottery number, ranging from 0 to 9. After all subjects in a session finish the questionnaire, the experimenter invites one of the subjects to draw a ten-sided die in front of all subjects in the session. Subjects who have a lottery number that equals the die roll get the earnings from decisions. One decision is randomly selected from the 42 decisions in the two parts as the *decision that counts*. If the decision that counts is from Part I, the allocation in that decision will be realized as the earnings from decisions. If the decision that counts is from Part II, the allocation will be realized and the subject will also receive a €5 participation fee on the additional date in Part II; hence a subject will receive three participation fees if a decision in Part II is realized. All the rules above were articulated in the instruction, and the instructions were always read aloud before either part of the experiment.

The earnings were paid by bank transfer to subjects' checking accounts. We made orders of transfers soon after the experiment and sent reminder emails with information about the incoming amounts on the experimental day and on all the payment dates. Given the reliability of the banking service, subjects can expect to receive all delayed payments exactly on the appropriate payment dates, while some of the present payments might be received one day after the experimental day due to the inter-bank processing.

We believe the payment tool we used was as good as cash in terms of liquidity. Checking accounts are used in private transactions such as paying for rents. Checking accounts are also linked to debit cards. In the Netherlands, debit cards are widely used for daily transactions in almost all kinds of stores including supermarkets, university restaurants and bookstores without any transaction fees. We held a survey about subjects' use of debit cards in the questionnaire. The responses show that bank transfers give high liquidity to the rewards, so that no isolation effect should be expected due to the payment method.[5]

[5] 84.7% of the subjects pay at least 50% of their expenditure in general by debit card, while 91.1% pay at least 30% of their expenditure in general by debit card. Among those who pay less than 30% of their expenditure in general by debit card, 61.1% pay at least 30% of their expenditure in university restaurants

2.1.4 *Transaction Costs and Credibility of Payments*

For our experiment, it is extremely important to equalize the transaction costs and the trustworthiness of the payments across periods because a difference in the transaction costs over the two periods can be a confounding factor of the magnitude effect.

Several facilities were employed in order to equalize the transaction costs across periods and to increase the credibility of the payments. The transaction costs include the costs to collect rewards, to confirm that the rewards have been received with correct amounts, and to remember the earnings so that they can be consumed on the expected dates.

First, we sent reminder emails with information about the incoming amounts on the experimental day and on all the payment dates. Subjects knew this from the instruction, so they did not need to worry about forgetting the earnings on the payment dates, a situation in which the expected marginal utility of the delayed rewards might be lowered.

Second, as Andreoni and Sprenger (2012) did, we delivered our business card and told the subjects to contact us immediately in case they would not receive a payment on time. It increased the credibility of payments and meanwhile served as a reminder of the payments.

Third, we asked subjects to fill in a payment reminder card with the amounts of their rewards on the corresponding dates just after their earnings were displayed. This served as a second reminder in case they forget to check emails.

In sum, the characteristics that one will receive a participation fee on each payment date and that all payments will be received by bank transfer help equalize the transaction costs of receiving payments on all dates. At the same time, the business cards, the payments reminder cards and the reminder emails reduced the risk of forgetting the rewards. The business cards also lowered the perceived

---

or in supermarkets by debit card. Among the remaining seven subjects, four withdraw cash at least 3 times per month.

default risks. Even though the risk might still be perceived by some subjects, it should be equal across periods since the payment tools and all auxiliary facilities were the same.

2.1.5 Sample

Our experiment was conducted at the CentERlab, Tilburg University in September of 2014.[6] 203 students of the university participated in one of the 11 sessions, 94 in the Present Group and 109 in the Delayed Group. Each subject made 42 decisions. One session took one hour and ten minutes on average. 22 subjects got the earnings from decisions, which averaged €69.16. The overall average earning was €17.49.

2.2 Hypotheses

The focus of this paper is on whether there is a magnitude effect on time preference, i.e. if people make intertemporal choices differently when the stakes vary. However, the definition of the magnitude effect still needs to be clarified.

In single-reward tasks, subjects reveal an indifference relation between a smaller sooner reward and a larger later reward, i.e. $(m_t, t) \sim (m_{t+\tau}, t+\tau)$. A monetary discount rate is then defined as $d_m \equiv \left(\frac{m_{t+\tau}}{m_t}\right)^{\frac{1}{\tau}} - 1$ where m_t and $m_{t+\tau}$ are revealed from the indifference relation. In this situation, the magnitude effect on time preference is defined on the monetary discount rate. A common result of such studies is that the monetary discount rate is decreasing in the stake, i.e. the monetary discount factor is increasing in the stake, which can be called as a positive magnitude effect on the monetary discount rate.

[6] The payment dates were in September, October and November. The fall semester in Tilburg University started from the end of August and ended in early December. Hence the payment dates were earlier than the final exam weeks and the Christmas vocation, which keeps our experiment from their probably large impacts on the subjects' demand of money.

In our intertemporal allocation task, the monetary discount rate cannot be defined, unless a subject always puts all the tokens onto one of the two dates. Therefore, the magnitude effect needs to be redefined. A natural way is to define it on the budget share: if the budget share allocated onto the later date ($\frac{n_{t+\tau}}{N}$, or equivalently $\frac{z_{t+\tau}}{m}$) is increasing in the size of the total budget (m), we call this a positive *magnitude effect on budget share*. The intuition is that when outcomes are scaled up, people are willing to postpone part of the sooner reward to the later date.[7] Formally, we test the following hypothesis:

Hypothesis 1 (magnitude effect on budget share): $\frac{z_{t+\tau}}{m}$ *is increasing in* m.

This hypothesis can be tested without assuming a specific model.

We are also interested in the relationship between the present bias and the magnitude effect. Benhabib et al. (2010) suggest that a fixed cost of delaying rewards can account for the present bias and the magnitude effect on monetary discount rates simultaneously, since the fixed cost induces the decision weight of future rewards to change disproportionately with delay and with the size of rewards. We would like to know if this cost is incurred only when a present reward is delayed or if it also applies to delaying a future reward. In a broader sense, we test whether the factors that drive the present bias (of which a fixed cost of delaying present rewards is an example) can account for the magnitude effect. If so, we should observe a magnitude effect in the Present Group, but not in the Delayed Group. If we do not observe a magnitude effect, or if we observe that the magnitude effects are of the same size in both groups, then it implies that the factors which drive the present bias cannot fully explain the magnitude effect. Thus, we establish our second hypothesis.

[7] The magnitude effect on the budget share requires the overall utility function not to be homogeneous. It can be with a stationary period utility function and a magnitude-independent discount rate. One example is the preference represented by $z_t + z_t^{0.5} + 0.9(z_{t+\tau} + z_{t+\tau}^{0.5})$, where z_t is a sooner reward and $z_{t+\tau}$ is a later reward.

Hypothesis 2 (no present reward, no magnitude effect): $\frac{z_{t+\tau}}{m}$ *does not change with* m *in the Delayed Group.*

Conditional on finding a positive magnitude effect, we wish to explore the channels of the magnitude effect. Given the evidence of time separability, we will estimate the parameters of preferences, with the assumption that subjects maximize a time separable utility function with CRRA atemporal utility functions and quasi-hyperbolic discounting, i.e. subjects maximize

$$U(z_t, z_{t+\tau}) = \frac{1}{\alpha}\delta^t(z_t + \omega)^\alpha + \beta\delta^{t+\tau}\frac{1}{\alpha}(z_{t+\tau} + \omega)^\alpha, \tag{2.1}$$

where β is the present bias parameter, δ is the daily discount factor, α is the exponent parameter. z_t and $z_{t+\tau}$ are the sooner reward and the later reward, respectively. ω is the background consumption mentally integrated with the experimental reward when the decision is made.

When the CRRA utility function is assumed, the elasticity of intertemporal substitution in consumption, $e_c \equiv -\frac{\ln\left(\frac{c_{t+\tau}}{c_t}\right)}{\ln\left(\frac{u'(c_{t+\tau})}{u'(c_t)}\right)}$, is equal to $\frac{1}{1-\alpha}$ (c_t and $c_{t+\tau}$ are the consumption on the sooner date and on the later date, respectively.). Thus, the exponent parameter, α, is a positive transformation of e_c. If $\alpha \to 1$, the atemporal utility function becomes linear, and the elasticity goes to infinity. In that case, subjects just go for the largest present value, and hence rewards are perfectly substitutable between dates. In case $\alpha \to -\infty$, the atemporal utility function is Leontief, and the elasticity goes to zero. In that case, subjects always divide the total budget into two equal amounts. In general, the larger the value of α, the more substitutable the subject considers the two rewards to be. Therefore, α is a measure of intertemporal substitutability.

It brings several advantages to assume such a model. First, the parameters in this model have important economic meanings. The discount factor determines the average choice across interest rates and hence measures the patience of the subject; if a subject is more patient, she will allocate more tokens to the later date for all interest rates. The intertemporal substitutability of consumption between different points in time relates to the dispersion of the choices across interest rates since it measures how sensitive the subject is to the interest rate. These behavioral measures are hard to estimate without assuming a model. Due to the non-negativity constraint, choices are censored at the corners if the preference parameters are extreme. As a result, directly measuring the average choice (as a measure of δ) and the dispersion of choices (as a measure of α) leads to biases. In contrast, the model we assume is tractable and easy to estimate. Moreover, the model is widely used in both theoretical and empirical applications. [8]

Given the model above, we test the following two hypotheses.

Hypothesis 3 (magnitude effect on discount factor): δ *is increasing in* m.

Hypothesis 4 (magnitude effect on intertemporal substitutability): α *is increasing in* m.

[8] To address the concern about misspecification, in Appendix 2.C, we check the robustness of our results by estimating a model with the Hyperbolic Absolute Risk Aversion (HARA) utility function and quasi-hyperbolic discounting. The HARA utility function is more flexible in the sense that it allows the atemporal utility function to be Increasing Relative Risk Aversion, Constant Relative Risk Aversion or Decreasing Relative Risk Aversion. This kind of flexibility is especially important when the magnitude is varied in the experiment. The results are the same.

2.3 Overall Effects

2.3.1 Magnitude Effect on Budget Share

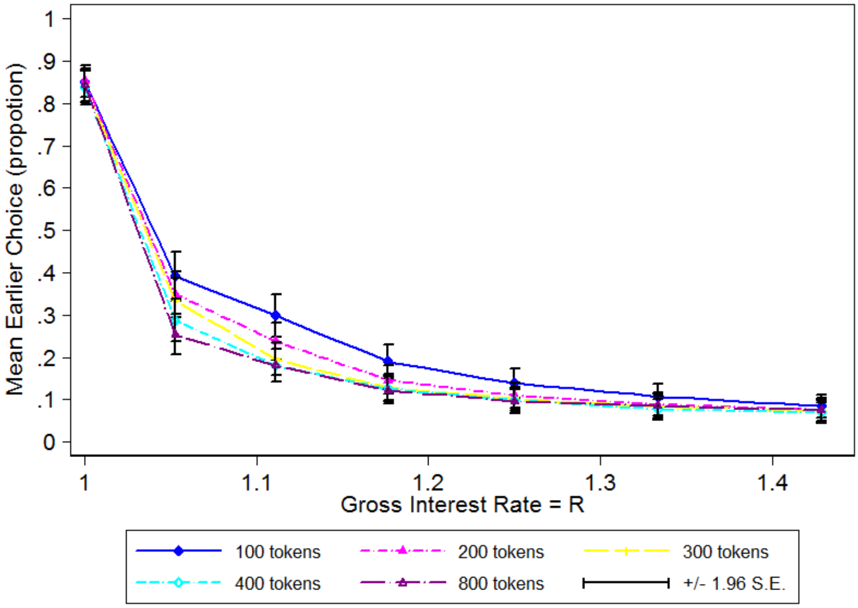


Figure 2.2: Mean Budget Share on the Sooner Date in Part I

In Figure 2.2 we plot the mean budget share allocated to the sooner date against the gross interest rate, R, of each CTB decision in Part I.[9] We plot separate points for the five magnitudes ($m =$ €20, €40, €60, €80, €160). The budget share allocated to the sooner date declines with the magnitude.

The difference seems to be larger when the interest rate is smaller but still positive. This is mainly due to censoring. When the interest rate is zero ($R = 1$) or highest ($R = 1.43$), most choices are at the corners for both smaller and larger magnitudes.

[9] In our data, 72% of the choices are at corners, and 38% of our subjects only make corner choices. This is very close to the 70% and 37%, respectively, in Andreoni and Sprenger (2012). The relationships between the budget shares and the interest rates are also similar.

Table 2.2: Mean Differences in Budget Shares on the Later Date between Magnitudes

	R						
	1	1.0526	1.1111	1.1765	1.25	1.3333	1.4286
€20 ~ €40	-0.0014	0.0417	0.0613	0.0438	0.0286	0.0186	0.0096
€40 ~ €60	0.0132	0.0157	0.0421	0.0179	0.0078	0.0053	0.0015
€60 ~ €80	0.0027	0.0463	0.0142	0.0024	0.0041	0.0067	0.0042
€80 ~ €160	-0.0066	0.0335	0.0008	0.0039	0.0020	-0.0076	-0.0058

Notes: Mean differences in the budget shares between two consecutive magnitudes given a gross interest rate, R.

To judge whether there is a significant magnitude effect, we perform Hotelling's T-squared tests on the mean differences in budget shares between magnitudes, taking seven choices with the same interest rate as a vector (see Table 2.3).[10] The null hypothesis is that the means of choices are the same across magnitudes, taking into account the correlation within subject. This class of tests makes sense because individual heterogeneity may have made different subjects reveal magnitude effects on tasks with different interest rates (e.g. Subject 1 on Interest Rate 1 while Subject 2 on Interest Rate 2), so that the magnitude effects on all choices would be jointly significant, but the effect on choices with any single interest rate might not be significant. The results show that the magnitude effect is significant between the magnitudes of €20 and €40 and between any two non-adjacent magnitudes. These results support Hypothesis 1, which states that a larger share of the budget is allocated to the later date when the size of the budget increases.[11]

[10] Hotelling's T-squared test is asymptotically nonparametric, so it can be applied to a large sample in nonnormal cases. We also perform a multivariate signed-rank test (Oja and Randles, 2004) and the results are basically the same: the magnitude effects are significant between the magnitudes of €20 and €40 and between any two non-adjacent magnitudes at least at the 10% level.

[11] In Table 2.3, statistics are reported only for four pairs of non-adjacent magnitudes. The mean differences for the other three pairs of non-adjacent magnitudes (in the two groups separately and in total) are also significant: seven out of the nine differences are significant at the 1% level, while the other two at the 5% level.

The results also show that the differences are insignificant between adjacent magnitudes larger than €20. Since the allocation is monotonic in the magnitude and the differences are significant between non-adjacent magnitudes, the insignificance suggests that the magnitude effect is largest when comparing the smallest magnitudes (€20 and €40), and becomes smaller for larger magnitudes. The pattern is consistent with the fact that Andersen et al. (2013) only found a small magnitude effect when they elicited time preferences using very high stakes.[12]

[12] They compare magnitudes of 1,500 Danish kroner and 3,000 Danish kroner, which is roughly equivalent to €200 and €400, respectively.

Table 2.3: Multivariate Mean Difference Tests between Magnitudes

		€20 ~ €40	€40 ~ €60	€60 ~ €80	€80 ~ €160	€20 ~ €60	€40 ~ €80	€60 ~ €160
Total	F-statistic	3.3184***	1.5477	1.5761	1.7104	5.1162***	3.9336***	3.0328***
	p-value	0.0023	0.1533	0.1444	0.1084	0.0000	0.0005	0.0047
Present Group	F-statistic	2.4091**	1.2374	1.6696	1.3919	3.2252***	2.9725***	2.0495*
	p-value	0.0266	0.2913	0.1270	0.2190	0.0044	0.0076	0.0577
Delayed Group	F-statistic	2.4388**	1.3650	1.0096	1.3659	2.8009**	2.1048**	2.0595*
	p-value	0.0237	0.2282	0.4290	0.2278	0.0104	0.0495	0.0547

Notes: Hotelling's T-squared tests on the mean differences in the budget shares between two magnitudes for all gross interest rates. 203 sets of observations for each magnitude. The degrees of freedom of the F-statistics are (7,196) in total, (7, 87) in the Present Group and (7, 102) in the Delayed Group. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

The results are robust against multiple hypotheses testing problem, since performing a Holm-Bonferroni correction on any family of four/three hypotheses does not change the significance at the 10 percent level.

2.3.2 *Present Bias*

Table 2.3 shows the results of the Hotelling's T-squared tests for the Present Group and for the Delayed Group separately. We find significant magnitude effects in both groups. This implies that the presence of an immediate reward is not a necessary condition for the magnitude effect. In other words, the factor that drives the present bias is unlikely to be the driver of the magnitude effect. Thereby, we reject Hypothesis 2.

We plot separate graphs for the two groups in Figure 2.3. Subjects in the Delayed Group seem to be slightly more patient than those in the Present Group. However, when we perform the Hotelling's T-squared test on all the 35 decisions in Part I between groups, the null hypothesis that the two groups have the same mean responses is not rejected. The p-value is 0.2424 when the degree of freedom is (35, 167). Thus, we find no evidence of present bias.[13]

Our result on the present bias and the magnitude effect is consistent with Sutter et al. (2013); they use binary choice lists to measure adolescents' time preferences and also find evidence of the magnitude effect but no evidence of the present bias.

Even though there might be a present bias which is not captured by our design due to lack of real immediate rewards, our results still have two implications for the magnitude effects. First, since a magnitude effect is present when the present bias is absent, our results imply that the factors which drive the present bias cannot fully account for the magnitude effect. Second, if it is a mental cost of delaying rewards that drives the magnitude effect (Benhabib and Bisin, 2005; Fudenberg and Levine, 2006), our results suggest that an equal size of mental cost

[13] The present bias here refers to non-stationarity of preferences according to the categorization of Halevy (2015). Our finding does not necessarily contradict the stylized fact that the discount rate is decreasing in the time distance between the sooner reward and the later reward, as in Benhabib et al. (2010). That stylized fact and stationarity can hold simultaneously if there is subadditivity in discounting (see Read, 2001).

is incurred when one postpones a future reward compared to when one postpones a present reward.

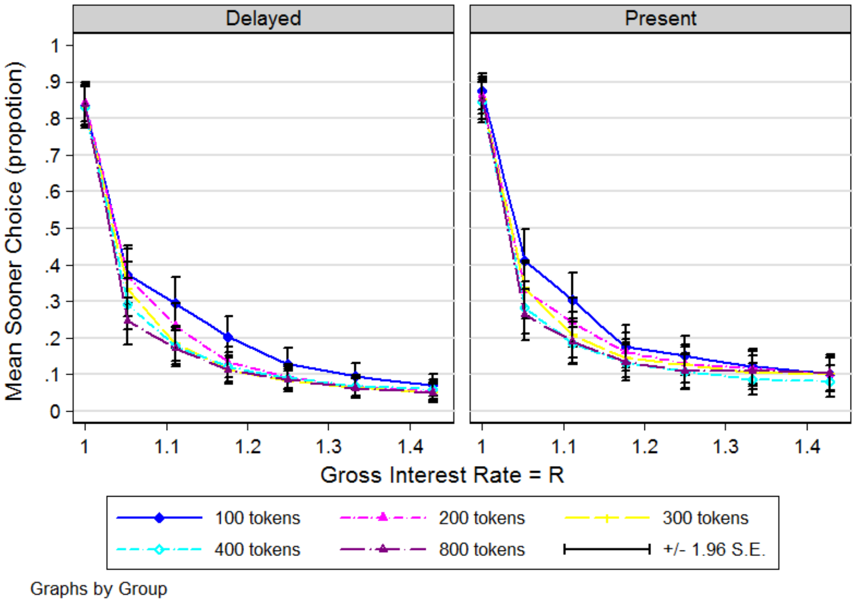


Figure 2.3: Mean Budget Shares in the Present Group and in the Delayed Group

2.3.3 *Time Separability*

The outcomes show that Part II is a valid test of time separability, since most subjects chose 200 tokens for the additional date in Part II. Only 39 out of 658 decisions from the Present Group and two out of 763 decisions from the Delayed Group were different from 200 tokens. Those involved eight subjects in the Present Group and one subject in the Delayed Group.

After removing those decisions, we compare the choices with the magnitude of €40 between Part I and Part II, separately for each group. Table 2.4 shows that the Hotelling's T-squared tests fail to reject the null hypothesis that responses to the two parts have the same means. Those results support time separability, which will be used in the following section.

Table 2.4: Multivariate Mean Difference Tests between Part I and Part II

Subsample	Present	Delayed	Total
F-statistic	1.5560	1.4192	1.0979
Degree of freedom	7, 79	7, 101	7, 187
p-value	0.1609	0.2058	0.3662

Notes: Hotelling's T-squared tests on the mean differences in the budget shares in the decisions with the magnitude of €40 between Part I and Part II. Subjects who chose a different number from 200 tokens for the additional date in Part II such that their choices were not comparable between the two parts have been removed from the sample. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

2.4 Channels

In order to disentangle the magnitude effect into two channels, we perform parametric estimations both at the aggregate level and at the individual level. We then test if the preference parameters change with the magnitude of the total budget.

2.4.1 Aggregate-Level Estimation

2.4.1.1 Estimation strategy

In our main specification, we assume a CRRA atemporal utility function as in equation (2.1). We set ω (background consumption) equal to the average response to the question about one's typical daily expenditure, €7.89, as Andreoni and Sprenger (2012) did in two of their specifications.[14]

[14] To fix the background consumption across subjects brings the advantage that all effects come from the variation in choices rather than also from the variation in the self-reported background consumptions, which may be noisy. We check the robustness by setting ω as individual background consumption, and average/individual background consumption combined with the participation fee (See Appendix 2.B). The results are basically the same.

Given the intertemporal utility function, solving the optimization problem yields the tangency condition

$$\frac{z_t+\omega}{z_{t+\tau}+\omega}=\begin{cases}(\beta\delta^{\tau}R)^{\frac{1}{\alpha-1}}, & \text{if } t=0\\ (\delta^{\tau}R)^{\frac{1}{\alpha-1}}, & \text{if } t>0\end{cases}.$$

Taking logs gives a linear equation

$$\ln\left(\frac{z_t+\omega}{z_{t+\tau}+\omega}\right)=\left(\frac{\ln\beta}{\alpha-1}\right)\cdot 1_{t=0}+\left(\frac{\ln\delta^{\tau}}{\alpha-1}\right)+\left(\frac{1}{\alpha-1}\right)\cdot\ln R$$

where $1_{t=0}$ is the indicator for the Present Group.

The parameters to be estimated are the present bias parameter, β, the discount factor, δ, and the CRRA curvature parameter, α. The present bias parameter is identified by the differences in allocation between the Present Group and the Delayed Group. If there is a present bias, subjects in the Present Group will allocate more tokens to the sooner date than those in the Delayed Group. The discount factor is identified by one's average choice across different experimental interest rates. A more patient subject will allocate more tokens to the later date in all decisions. The curvature parameter is identified by the dispersion of one's choices across interest rates. Those who consider rewards highly substitutable over time are likely to make corner choices in all decisions, while those with lower elasticity of intertemporal substitution will make choices closer to equal splits.

Following the practice in previous studies (Andreoni and Sprenger, 2012; and Augenblick, Niederle and Sprenger, 2015), we assume a normally distributed error term additive to the log allocation ratio and take censoring into consideration, then we yield the two-limit Tobit model:

$$l^*_{i,j,k}\equiv\ln\left(\frac{z^*_{t;i,j,k}+\omega}{z^*_{t+\tau;i,j,k}+\omega}\right)$$
$$=\left(\frac{\ln\beta}{\alpha-1}\right)\cdot 1_{t=0;i}+\left(\frac{\ln\delta^{\tau}}{\alpha-1}\right)+\left(\frac{1}{\alpha-1}\right)\ln R_j+\epsilon_{i,j,k},\epsilon_{i,j,k}\sim N(0,\sigma_k)$$

$$l_{i,j,k} = \begin{cases} \ln \dfrac{\omega}{m_k + \omega}, & \text{if } l^*_{i,j,k} \leq \ln \dfrac{\omega}{m_k + \omega} \\ l^*_{i,j,k}, & \text{if } \ln \dfrac{\omega}{m_k + \omega} < l^*_{i,j,k} < \ln \dfrac{\frac{m_k}{R_j} + \omega}{\omega} \\ \ln \dfrac{\frac{m_k}{R_j} + \omega}{\omega}, & \text{if } l^*_{i,j,k} \geq \ln \dfrac{\frac{m_k}{R_j} + \omega}{\omega} \end{cases}$$

where $i = 1, \dots, 203$ denotes Subject i, $j = 1, \dots, 7$ denotes Interest rate j, and $k = 1, \dots, 5$ denotes Magnitude k. The error term is allowed to vary across magnitudes since giving a larger number of tokens might induce a larger noise, which might be a competing explanation of a larger sensitivity to the interest rate.

The model is estimated by the quasi-maximum-likelihood method: when performing the estimation, the error term, $\boldsymbol{\epsilon}$, is assumed to be i.i.d., while in computing the standard errors, the error term is assumed to be independent across subjects, but might be correlated within-subject. Estimates of the parameters can be recovered and standard errors can be inferred by the delta method.

Since we are interested in the magnitude effect, we also perform the estimation with interaction terms of the parameters and the magnitude dummies. Thus, tests can be performed on the differences between the parameters for different magnitudes.

In Appendix 2.C, we assume another specification, in which the utility function is Hyperbolic Absolute Risk Aversion (HARA). In that specification the background consumption, ω, is also a parameter to be estimated. In this way, we address the concern that the average self-reported background consumption may not match the true background consumption integrated with the experimental rewards in decision making, or the Relative Risk Aversion of the utility function may not be constant (i.e. the CRRA utility function is misspecified). The results are basically the same.

2.4.1.2 Results

Table 2.5 reports the magnitude-invariant estimates and the magnitude-specific estimates of the parameters, respectively. A salient feature is that none of the estimates of β is significantly different from 1, implying no evidence of present bias, which is consistent with our finding in the model-free analysis. The annual discount rate for all magnitudes is 52.7%, which is in the range found by previous studies. The CRRA curvature parameters are always significantly smaller than 1, implying that the subjects on average consider the monetary rewards received on different dates imperfectly substitutable, which is also consistent with other studies (e.g. Andreoni and Sprenger, 2012; Andreoni, Kuhn and Sprenger, 2013; Cheung, 2015; and Augenblick, Niederle and Sprenger, 2015).

Most importantly, both the discount factor and the CRRA curvature are increasing in the magnitude. To judge if these magnitude effects are significant, Table 2.6 presents Wald tests over the differences of parameters between magnitudes.[15] We find significant magnitude effects both on the discount factor, δ, and on the exponent parameter, α, which is a positive transformation of the elasticity of intertemporal substitution. The discount factor is increasing in the magnitude, which is consistent with previous studies. The elasticity of intertemporal substitution is increasing in the magnitude, meaning that the rewards on the two dates are more substitutable to the subjects when the subjects face a larger total budget. This results in choices closer to the two corners (to which corner depends on whether $\delta R > 1$). Thereby, we verify Hypothesis 3 and Hypothesis 4.

[15] For the other three pairs of non-adjacent magnitudes: the differences in β are not significant, while the differences in δ^{τ} and in α are all significant at the 1% level.

Table 2.5: Discounting and Curvature Parameter Estimates in the Aggregate-Level Estimation with the CRRA Specification

Model:	Tobit			Tobit		
Magnitude:	All	€20	€40	€60	€80	€160
Present bias: $\hat{\beta}$	0.989	0.989	0.989	0.986	0.997	0.986
	(0.018)	(0.021)	(0.018)	(0.018)	(0.017)	(0.020)
Discount factor over four weeks: $\widehat{\delta^{\tau}}$	0.968	0.948	0.961	0.971	0.972	0.982
	(0.012)	(0.016)	(0.013)	(0.012)	(0.011)	(0.012)
CRRA curvature: $\hat{\alpha}$	0.955	0.928	0.947	0.952	0.958	0.968
	(0.004)	(0.007)	(0.005)	(0.005)	(0.004)	(0.003)
S.e. of the error term: $\hat{\sigma}$	3.699	2.294	2.986	3.369	3.857	5.314
	(0.281)	(0.200)	(0.245)	(0.269)	(0.307)	(0.454)
Log-likelihood	-13678.51			-13538.56		
Observations	7,105			7,105		
Uncensored	1,969			1,969		
Clusters	203			203		

Notes: Two-limit Tobit estimators. CRRA estimation with $\omega = 7.89$ (average reported background consumption). Column 1: assuming that parameters are invariant to magnitudes. Column 2-6: assuming that parameters vary with magnitudes. Clustered standard errors in parentheses. Log-likelihood has been corrected for the transformation of dependent variables. Standard errors calculated via the delta method.

Table 2.6: Estimates of Parameter Differences between Magnitudes in the CRRA Specification

Magnitude:	€40 - €20	€60 - €40	€80 - €60	€160 - €80	€60 - €20	€80 - €40	€160 - €60
Present bias: $\hat{\beta}$	-0.000	-0.003	0.011*	-0.011	-0.003	0.008	0.000
	(0.008)	(0.006)	(0.006)	(0.009)	(0.010)	(0.008)	(0.009)
Discount factor over four weeks: $\widehat{\delta^{\tau}}$	0.014**	0.010*	0.001	0.010**	0.024***	0.011*	0.011*
	(0.006)	(0.005)	(0.005)	(0.004)	(0.008)	(0.006)	(0.006)
CRRA curvature: $\hat{\alpha}$	0.018***	0.006***	0.006***	0.011***	0.024***	0.011***	0.016***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)
S.e. of the error term: $\hat{\sigma}$	0.692***	0.383***	0.487***	1.457***	1.075***	0.871***	1.944***
	(0.106)	(0.092)	(0.108)	(0.218)	(0.138)	(0.125)	(0.258)

Notes: Estimates of parameter differences are inferred from the Two-limit Tobit estimation by the delta method. The estimation assumes CRRA utility with ω =7.89. Separate parameters are estimated for each magnitude among €20, €40, €60, €80 and €160. There are 1,421 observations (203 clusters) for each magnitude. Clustered standard errors in parentheses. Standard errors calculated via the delta method. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

To get an idea about the relative importance of the two channels of the magnitude effect, we use the estimates above to predict choices in the 35 questions for both the Present Group and the Delayed Group. Table 2.7 presents the marginal effects of allowing one parameter to vary with the magnitude: in each row, we allow only one parameter, either δ or α, to vary with the magnitude of the decisions (as indicated by the column title), but fix the other two parameters at the value estimated from the magnitude of €20. Each number in a cell is the total change (in unit of $\frac{N_k}{100}$, the percentage of the total budget) in the seven decisions with the corresponding magnitude. The results show that the marginal effect of allowing α to vary with the magnitude is at least as large as the marginal effect of allowing δ to vary. This suggests that the magnitude effect on the elasticity of intertemporal substitution is at least as important as the magnitude effect on the discount rate.

Table 2.7: Marginal Effects of Allowing a Parameter to Vary with Magnitudes in the CRRA Specification

	Magnitude:	€40	€60	€80	€160
Parameter values used in prediction:	$\beta_1, \delta_k, \alpha_1$ (Delayed):	21.3	34.7	36.6	48.6
	$\beta_1, \delta_k, \alpha_1$ (Present):	21.8	36.5	38.1	50.7
	$\beta_1, \delta_1, \alpha_k$ (Delayed):	24.8	33.9	44.4	65.5
	$\beta_1, \delta_1, \alpha_k$ (Present):	27.7	38.0	49.2	72.8

Notes: The changes in choices predicted by the CRRA Tobit model using the parameter values indicated by the row title compared with $(\beta_1, \delta_1, \alpha_1)$, for the two groups separately. k in the row titles stands for the magnitude in the column title. For instance, the first cell in the first row is the difference between the choices made in the seven decisions with the magnitude of €40 predicted by the model with parameter values $(\beta_1, \delta_2, \alpha_1)$ and those predicted by the model with parameter values $(\beta_1, \delta_1, \alpha_1)$. In other words, it is the marginal effect of allowing δ to vary with the magnitude from €20 to €40. The unit is 1 percent of the total budget.

2.4.2 *Individual-Level Estimation*

The aggregate-level estimation provides evidence of positive magnitude effects on the discount factor and on intertemporal substitutability. One may wonder whether these results also hold at the individual level. Indeed, we find a huge individual heterogeneity in choices. One concern is that, when testing the magnitude effect on the aggregate preferences, there might be a bias resulting from forcing all subjects to have the same preferences and the same distribution of noise. To deal with this concern, we also perform individual-level estimation and tests.

2.4.2.1 Estimation and testing procedure

We keep all the assumptions that underlie equation (2.1) except for β since it is not identified in individual-level estimations. We estimate the discount factor (δ) and the intertemporal substitutability (α) for each combination of subject and stake, and then test if the two parameters are increasing in the magnitude within-subject.

One important difference from the aggregate-level estimation is that there might be an under-identification problem when a subject made no or only one interior choice under a stake. Actually, there are 627 out of 1015 (62%) combinations of subjects and stakes suffering from such a problem. We thereby adopt a conservative way to test the magnitude effect. First, we yield point estimates of δ and α if possible. Whenever there is an under-identification problem, we remove the error term from (2.1) and then infer the intervals of δ and α that can generate the observations. Second, we perform a one-tailed sign test on the two parameters, respectively, with the null hypotheses that they do not change with the magnitude. The sign test only requires that the distribution of a parameter does not differ between magnitudes, while it allows the distribution to be different across subjects. For a comparison between a point estimate and an interval estimate, we recognize a difference only if the point is not in the interior of the interval. For a comparison between two interval estimates, we recognize a difference if the two intervals do not overlap.

2.4.2.2 Results

Table 2.8 shows the results of the tests at the individual level. We reject the null hypotheses of no magnitude effect on the two parameters, in favor of positive magnitude effects. This shows that the two channels of the magnitude effect on intertemporal choices are robust against individual heterogeneity.

Table 2.8: Sign Tests on Preference Parameters between Magnitudes

		€20 ~ €40	€40 ~ €60	€60 ~ €80	€80 ~ €160	€20 ~ €60	€40 ~ €80	€60 ~ €160
Discount factor over four weeks: $\widehat{\delta^{\tau}}$	#increase /unchanged /decrease	31% / 48% / 21%	31% / 51% / 19%	26% / 56% / 18%	23% / 58% / 19%	36% / 48% / 16%	34% / 50% / 16%	29% / 54% / 17%
	p-value	0.0255**	0.0107**	0.0687*	0.2253	0.0001***	0.0003***	0.0086***
CRRA curvature: $\hat{\alpha}$	#increase /unchanged /decrease	30% / 61% / 9%	28% / 60% / 12%	24% / 63% / 13%	27% / 64% / 9%	31% / 62% / 8%	29% / 61% / 10%	29% / 63% / 9%
	p-value	0.0000***	0.0002***	0.0101**	0.0000***	0.0000***	0.0000***	0.0000***

Notes: Right-tailed sign tests on the differences of parameters between two magnitudes. 203 sets of observations for each magnitude. $\omega = 7.888$. When under-identification occurs, interval estimates are yielded for the two preference parameters. A point estimate and an interval estimate are considered as different only if the point is not in the interior of the interval. Two interval estimates are considered as different only if their intersection is empty or a singleton. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

2.5 Interpretations

The results above imply that when an average subject faces a larger budget in an intertemporal allocation task, she behaves more patiently, but also she regards rewards to be more substitutable between dates.

2.5.1 Relation with the Magnitude Effect on Risk Aversion

According to the Discounted Expected Utility (DEU) theory, the risk attitude and the elasticity of intertemporal substitution are represented by the same parameter, since risk aversion and imperfect fungibility both originate from diminishing marginal utility. Therefore, one may wonder whether the magnitude effect on intertemporal substitutability is the same as the magnitude effect on risk attitudes.

We find evidence against this equivalence. Holt and Laury (2002) investigated the magnitude effect on risk attitudes with Multiple Price List (MPL) questions. They found a significant, positive magnitude effect: when faced with a larger magnitude, people appear to be more risk averse in terms of the relative risk aversion. This is in the opposite direction as the effect we find. Their finding suggests an increase in the concavity as the magnitude increases while ours shows a movement towards linearity. This contradiction suggests that the magnitude effect on relative risk aversion is not driving the magnitude effect on intertemporal substitutability.

Some other studies also suggest that risk aversion and intertemporal substitutability should be separated. Andreoni and Sprenger (2012) found no significant correlation at the individual level between the curvature estimated by the CTB method and the risk attitude elicited by the MPL method. Abdellaoui et al. (2013), Miao and Zhong (2015) and Cheung (2015) also found that the utility curvature elicited from intertemporal tasks is different from that elicited from tasks with risk. We provide evidence from a different perspective: while the previous studies showed that the degrees of concavity are different for the two kinds of utility functions, we show that the degrees of concavity change in opposite directions when the stake is varied.

This finding has implications for both theories and experimental methods. First, it lends support to the theories which separate intertemporal substitutability from risk aversion, such as Epstein and Zin (1989). Second, it casts doubt on the use of a risk-elicitation task to correct for the curvature when eliciting time preferences.

2.5.2 Relation with Borrowing Constraints

In theory, a binding borrowing constraint can lead to a magnitude effect on the monetary discount rate in a single-reward task if the background consumption is expected to grow over time, as shown by Epper (2015). However, Meier and Sprenger (2010) found that experimentally elicited long-run discount rates are uncorrelated with credit constraints, suggesting that on average, whether the borrowing constraint is binding does not affect intertemporal choices in experiments.

Moreover, given the fact that subjects may have savings which provide limited liquidity, the fraction of subjects whose borrowing constraints are binding is increasing in the stake. For this reason, if the borrowing constraint is a main issue, we should observe that the intertemporal substitutability is decreasing in the stake, which is inconsistent with our results. Therefore, we believe that a binding borrowing constraint is not the main driver of our results.

2.5.3 Relation with Existing Theories

We discuss the implications of our empirical findings for some theories that may explain the magnitude effect on intertemporal choices.

One model that can account for the magnitude effect on the discount factor was proposed by Benhabib et al. (2010). They developed a model with a fixed cost of delaying rewards. The idea is that whenever a delayed reward is chosen, a fixed cost is incurred, so that as the stake increases, the cost becomes relatively less important and hence the subject appears more patient.

Noor (2011) proposed a model of magnitude-dependent discounting, which leads to similar predictions. In his model, the discount function is increasing in the utility at the later period. As the stake gets larger, the discount function converges to 1. [16]

One theory that can explain the magnitude effect on intertemporal substitutability is an extended version of the dual-self bank-nightclub model of Fudenberg and Levine (2006). In the original model, the agent chooses the amount of pocket cash when no temptation is present, and then she chooses the amount of consumption when a windfall is available and temptation plays a role. The strategy for utility maximization is to spend all of the windfall when it is small but try to save some money out of the windfall when it is large. A small windfall is not integrated with the lifelong wealth, because the agent does not bother to perform self-control, but it is worth controlling oneself when the windfall is large. As a result, the utility function for windfalls is much more concave when the size of the windfall is below a certain threshold than when it is above the threshold.

The model can explain a magnitude effect on intertemporal substitutability if we impose the assumption that an agent who anticipates a reward in the future does not immediately adjust her cash allocation plan. Instead, she keeps the anticipated reward in the mental account of windfalls until it is received and part

[16] Both models can explain the magnitude effect on the monetary discount rate in single-reward tasks. When they are applied to the intertemporal allocation tasks, both of them predict a jump from the sooner corner to an interior point or to the later corner in every wealth expansion path. Unfortunately, we are not able to confirm nor reject the existence of such jumps; since we have only five points on each wealth expansion path, it is hard to distinguish jumps from curves.

of it is consumed. Only after the remainder is moved into the mental account of savings does she reschedule her future consumption.

When this assumption is used, the model predicts that a subject will tend to make interior choices when the budget is small, i.e., below the threshold induced by the self-control costs. Since the utility function for windfalls is very concave the subject balances extra consumption on the sooner date and on the later date. As the budget increases above the threshold, the subject will want to save part of it for consumption smoothing. Since the utility function for savings is much less concave (close to linear) these savings will be allocated fully to either the sooner date (when the interest rate is small) or the later date (when the interest rate is large). Hence, as the budget increases the intertemporal substitutability increases and it will appear as if the utility function has become less concave (see Appendix 2.D for a simulation).

Another model that can explain the magnitude effect on intertemporal substitutability is the mental zooming theory proposed by Holden (2014). The theory presumes that people integrate more background consumption with the experimental reward as the size of the reward increases. If the budget increases, individuals 'zoom out' as it were, and take a broader perspective in the decision problem. One reason may be that individuals are likely to divide and use up a bigger windfall over a longer time period. Based on the data collected from his field experiment with Malawian peasants, Holden showed that the magnitude effect on time preferences in single-reward tasks would disappear if the unobserved background consumption is assumed to be an increasing function of the stake.

In intertemporal allocation tasks, the increasing background consumption can generate a magnitude effect on intertemporal substitutability. To see why, denote the observed elasticity of intertemporal substitution in experimental rewards by e_z. The relationship between e_z and preference parameters is

$$e_z = \frac{1}{1-\alpha} \cdot \frac{\log\left(\frac{z_{t+\tau}}{z_t}\right)}{\log\left(\frac{z_{t+\tau}+\omega}{z_t+\omega}\right)}.$$

Since e_z is increasing in both α and ω, an increase in α and an increase in ω are competing explanations for the magnitude effect on intertemporal substitutability. If subjects take into account more background consumption as the total budget increases, we would observe a greater sensitivity to the interest rate, i.e. a greater e_z. When we assume a fixed background consumption, however, the pattern will be attributed to a magnitude effect on α.

Both the mental-accounting Fudenberg-Levine model and the mental zooming theory point to partial integration with lifelong wealth, which seems to be an important mechanism of the magnitude effect on intertemporal substitutability between rewards. Andersen et al. (2012) showed empirically that subjects only partially integrate experimental rewards with wealth in risk preference tasks. While they provide evidence of partial asset integration by exploiting variation in personal wealth, our results suggest that the degree of asset integration is increasing in the stake by providing within-subject evidence.

None of the current models can explain both a magnitude effect on the discount factor and a magnitude effect on the intertemporal substitutability. Of course, the two channels can be explained by a mode-switching model in which individuals are assumed to have different preferences for different stakes. However, a truly unified explanation is still lacking.

2.6 Conclusion

Our study investigates the magnitude effect on intertemporal choices in a recently-introduced task, the intertemporal allocation task. After adapting the concept for the new task, we verify the existence of a magnitude effect. The magnitude effect exists even when both rewards are only available in the future, implying that the factor which drives the present bias cannot fully account for the magnitude effect.

We then look deeper into the effect, by exploring the channels. The results underscore the importance of a dimension which is often overlooked, namely, the intertemporal substitutability. We find evidence that both the discount factor and the intertemporal substitutability change with the magnitude of rewards.

Some existing theories may provide explanations for one of the two channels. A cost-of-delay model (Benhabib et al., 2010) or a magnitude-dependent discounting model (Noor, 2011) can account for a magnitude effect on the discount factor.[17] Models which allow the degree of asset integration (mental accounting) to vary with the size of the budget can explain a magnitude effect on intertemporal substitutability. However, a new theory would be needed to account for both channels simultaneously and in a unified way.

For the magnitude effect on intertemporal substitutability, existing theories tend to attribute it to the varying degree of asset integration, however, sharper tests are needed to check the conjecture and to explore specific factors. One possible way would be to restrict the dates on which rewards can be consumed and then to check if the restriction has an effect on intertemporal choices.

[17] Baucells and Heukamp (2012) extend the magnitude-dependent discounting model to allow risks.

APPENDIX TO CHAPTER 2

2.A Rationality of Subjects in the Convex Time Budget Method

The CTB method allows subjects to make interior choices, and hence makes it possible to measure discount rates and utility curvature simultaneously. However, Chakraborty et al. (2017) found that a proportion of subjects, especially those who make interior choices, violate wealth monotonicity in CTB tasks and that the magnitude of wealth monotonicity violations conditional on violating at least once are as large as that generated by uniform random choice, and hence questioned the rationality of subjects in making CTB decisions.

In this appendix, we follow Chakraborty et al. (2017) to examine price monotonicity and wealth monotonicity of our dataset. In specific, we look at fractions of monotonicity violations among all subjects and among subjects who make at least one interior choice, respectively. We also measure the average magnitude of wealth monotonicity violations for those who violate wealth monotonicity at least once, and we compare it with the distribution of the magnitude generated by uniform random choice.

Table A2.1 shows the rationality indices of the full sample and the subsamples as well as those generated by uniform random choice. The fractions of price monotonicity violations (2-3%) are less than the fractions of wealth monotonicity violations (10-20%). A possible reason is that prices are varied within decision forms but total budgets are varied across decision forms, and hence it is easier to make comparison across prices than across total budgets. Nevertheless, those fractions and the magnitudes of violations are much lower than those generated by uniform random choice, suggesting that our dataset and the subsample of subjects who make interior choices are highly informative.

Table A2.1: Rationality of Subjects Compared to Uniform Random Choice

	Fraction of price monotonicity violations	Average magnitude of price monotonicity violations (euros)	Fraction of wealth monotonicity violations	Average magnitude of wealth monotonicity violations (euros)
Full sample	0.022	0.13	0.101	0.82
Subjects who make at least one interior choice	0.037	0.23	0.155	1.03
Subjects who violate wealth monotonicity at least once	0.033	0.26	0.208	1.67
Mean of uniform random choice (Standard deviation)	0.429 (0.006)	8.12 (0.21)	0.297 (0.008)	2.69 (0.10)

Notes: The first three rows present fractions of price monotonicity violations and wealth monotonicity violations as well as average magnitudes of violations in terms of euros in the full sample, the subsample of subjects who make at least one interior choice, and the subsample of subjects who violate wealth monotonicity at least once. The last row presents the means and the standard deviations of the same four indices generated by uniform random choice. The means and standard errors are calculated by simulating 10,000 times.

2.B Parametric Analysis with Different Assumptions on Background Consumption

We check the sensitivity of the parameter estimates (Table A2.2) and the magnitude effects (Table A2.3) to alternative assumptions on the background consumption. The results show that the magnitude effects on the discount factor and on intertemporal substitutability are robust, except when ω is assumed to be very small. However, with a small background consumption, one should rarely make corner choices due to the motivation of consumption smoothing. But we do observe many corner choices in our sample. Hence, the case with a small ω is unlikely to be true.

Table A2.2: Background Consumption, Parameter Estimates and Likelihood

	Model:	Tobit			Tobit		
	Magnitude:	All	€20	€40	€60	€80	€160
$\omega = 7.89$	Present bias: $\hat{\beta}$	0.989	0.989	0.989	0.986	0.997	0.986
		(0.018)	(0.021)	(0.018)	(0.018)	(0.017)	(0.020)
	Discount factor over	0.968	0.948	0.961	0.971	0.972	0.982
	four weeks: $\widehat{\delta^{\tau}}$	(0.012)	(0.016)	(0.013)	(0.012)	(0.011)	(0.012)
	CRRA curvature: $\hat{\alpha}$	0.955	0.928	0.947	0.952	0.958	0.968
		(0.004)	(0.007)	(0.005)	(0.005)	(0.004)	(0.003)
	S.e of the error term:	3.699	2.294	2.986	3.369	3.857	5.314
	$\hat{\sigma}$	(0.281)	(0.200)	(0.245)	(0.269)	(0.307)	(0.454)
	Log-likelihood	-13678.51			-13538.56		
$\omega =$ $\max(\omega_i, 0.01)$	Present bias: $\hat{\beta}$	0.990	0.989	0.990	0.987	0.998	0.987
		(0.018)	(0.020)	(0.018)	(0.015)	(0.017)	(0.022)
	Discount factor over	0.968	0.948	0.962	0.971	0.972	0.982
	four weeks: $\widehat{\delta^{\tau}}$	(0.012)	(0.014)	(0.012)	(0.011)	(0.011)	(0.012)
	CRRA curvature: $\hat{\alpha}$	0.962	0.945	0.956	0.960	0.964	0.972
		(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.003)
	S.e of the error term:	4.422	3.048	3.716	4.102	4.570	6.115
	$\hat{\sigma}$	(0.358)	(0.300)	(0.315)	(0.328)	(0.370)	(0.552)
	Log-likelihood	-13920.57			-13815.64		
$\omega = 12.89$	Present bias: $\hat{\beta}$	0.989	0.989	0.989	0.985	0.997	0.986
		(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)
	Discount factor over	0.969	0.948	0.962	0.972	0.973	0.982
	four weeks: $\widehat{\delta^{\tau}}$	(0.012)	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)
	CRRA curvature: $\hat{\alpha}$	0.945	0.903	0.932	0.940	0.948	0.963
		(0.005)	(0.010)	(0.007)	(0.006)	(0.005)	(0.004)

	S.e of the error term:	3.013	1.697	2.329	2.698	3.144	4.490
	$\hat{\sigma}$	(0.229)	(0.147)	(0.187)	(0.214)	(0.262)	(0.384)
	Log-likelihood	-13623.90			-13443.15		
$\omega = \omega_i + 5$	Present bias: $\hat{\beta}$	0.990	0.990	0.990	0.987	0.998	0.987
		(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
	Discount factor over	0.968	0.948	0.962	0.971	0.972	0.982
	four weeks: $\widehat{\delta^{\tau}}$	(0.012)	(0.014)	(0.013)	(0.012)	(0.013)	(0.012)
	CRRA curvature: $\hat{\alpha}$	0.948	0.912	0.936	0.944	0.951	0.964
		(0.005)	(0.009)	(0.006)	(0.006)	(0.005)	(0.004)
	S.e of the error term:	3.213	1.870	2.529	2.899	3.355	4.430
	$\hat{\sigma}$	(0.246)	(0.163)	(0.205)	(0.231)	(0.275)	(0.398)
	Log-likelihood	-13693.32			-13527.03		
$\omega = 0.01$	Present bias: $\hat{\beta}$	0.990	0.990	0.990	0.987	0.998	0.986
		(0.019)	(0.020)	(0.019)	(0.018)	(0.020)	(0.022)
	Discount factor over	0.964	0.945	0.959	0.968	0.969	0.980
	four weeks: $\widehat{\delta^{\tau}}$	(0.013)	(0.015)	(0.013)	(0.010)	(0.013)	(0.011)
	CRRA curvature: $\hat{\alpha}$	0.989	0.989	0.989	0.989	0.989	0.989
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	S.e of the error term:	15.333	14.511	14.500	14.470	15.251	17.613
	$\hat{\sigma}$	(1.216)	(1.239)	(1.154)	(1.115)	(1.220)	(1.478)
	Log-likelihood	-15596.78			-15575.13		
	Observations	7,105			7,105		
	Uncensored	1,969			1,969		
	Clusters	203			203		

Notes: Two-limit Tobit estimators. Panel 1: $\omega =$ average reported background consumption. Panel 2: $\omega =$ individual reported background consumption (except for one subject, we replace the zero consumption with 0.01). Panel 3: $\omega =$ average reported background consumption plus the participation fee. Panel 4: $\omega =$ individual background consumption plus the participation fee. Panel 5: $\omega = 0.01$. Column 1: assuming that parameters are invariant to magnitudes. Column 2-6: assuming that parameters vary with magnitudes. Clustered standard errors in parentheses. Log-likelihood has been corrected for the transformation of dependent variables. Standard errors calculated via the delta method.

Table A2.3: Background Consumption and Magnitude Effects

	Magnitude:	€40 - €20	€60 - €40	€80 - €60	€160 - €80	€60 - €20	€80 - €40	€160 - €60
$\omega = 7.89$	Present bias: $\hat{\beta}$	-0.000	-0.003	0.011*	-0.011	-0.003	0.008	0.000
	Discount factor over four weeks: δ^{τ}	0.014**	0.010*	0.001	0.010**	0.024***	0.011*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.018***	0.006***	0.006***	0.011***	0.024***	0.011***	0.016***
	S.e of the error term: $\hat{\sigma}$	0.692***	0.383***	0.487***	1.457***	1.075***	0.871***	1.944***
$\omega = \max(\omega_i, 0.01)$	Present bias: $\hat{\beta}$	0.001	-0.003	0.011*	-0.011	-0.002	0.008	-0.000
	Discount factor over four weeks: δ^{τ}	0.013**	0.010**	0.001	0.010**	0.023***	0.010*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.011***	0.004**	0.004***	0.008***	0.015***	0.008***	0.012***
	S.e of the error term: $\hat{\sigma}$	0.668***	0.386***	0.468***	1.545***	1.054***	0.854***	2.013***
$\omega = 12.89$	Present bias: $\hat{\beta}$	-0.000	-0.003	0.011*	-0.011	-0.003	0.008	0.000
	Discount factor over four weeks: δ^{τ}	0.014***	0.010**	0.001	0.010**	0.024***	0.011*	0.011*

	CRRA curvature: $\hat{\alpha}$	0.028***	0.009***	0.008***	0.014***	0.037***	0.017***	0.022***
	S.e of the error term: $\hat{\sigma}$	0.632***	0.370***	0.446***	1.346***	1.002***	0.815***	1.792***
$\omega = \omega_i + 5$	Present bias: $\hat{\beta}$	0.001	-0.003	0.011*	-0.011	-0.002	0.008	-0.000
	Discount factor over four weeks: δ^{τ}	0.014***	0.010*	0.001	0.010**	0.023***	0.011*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.025***	0.008***	0.007***	0.013***	0.032***	0.015***	0.020***
	S.e of the error term: $\hat{\sigma}$	0.659***	0.371***	0.456***	1.374***	1.030***	0.827***	1.830***
$\omega = 0.01$	Present bias: $\hat{\beta}$	-0.000	-0.002	0.010	-0.011	-0.003	0.008	-0.001
	Discount factor over four weeks: δ^{τ}	0.014**	0.009	0.001	0.011**	0.023***	0.010	0.011*
	CRRA curvature: $\hat{\alpha}$	0.000	-0.000	0.001	0.001***	-0.000	0.000	0.002***
	S.e of the error term: $\hat{\sigma}$	-0.011	-0.030	0.781*	2.362***	-0.041	0.751	3.143***

Notes: Estimates of parameter differences are inferred from the Two-limit Tobit estimation by the delta method. Parameters are separately set for each magnitude among €20, €40, €60, €80 and €160. Panel 1: $\omega =$ average reported background consumption. Panel 2: $\omega =$ individual reported background consumption (except for one subject, we replace the zero consumption with 0.01). Panel 3: $\omega =$ average reported background consumption plus the participation fee. Panel 4: $\omega =$ individual background consumption plus the participation fee. Panel 5: $\omega = 0.01$. There are 1,421 observations (203 clusters) for each magnitude. Clustered standard errors in parentheses. Standard errors calculated via the delta method. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

2.C Parametric Analysis with Estimation of Background Consumption

The CRRA specification with exogenously set background consumption (ω) is simple and easy to estimate, however, one may suspect that the average self-reported background consumption does not match the true background consumption integrated with the experimental rewards in decision making, or the utility function is misspecified. In particular, if the utility function is Hyperbolic Absolute Risk Aversion (HARA), ω in (2.1) does not represent the background consumption but is a mixture of the background consumption and the HARA intercept parameter.

In order to meet the challenges above, we estimate ω instead of setting it exogenously. By doing this, we "let the data tell" what values are suitable for ω, and we can also identify the magnitude effects on ω.

2.C.1 Model

We assume a normally distributed error term at the choice level. The error term can be arisen either because of idiosyncratic shocks in preference or because of imprecision in placing the scrollbar.[18] In specific, a latent choice is

[18] In the model with a normally distributed error term additive to the log allocation ratio, the estimator of ω is nonlinear in the error term. Simulation shows that the estimator of ω is severely biased given our sample size, though it is asymptotically consistent. The model we assume here is the same as the one implicitly assumed by Andreoni and Sprenger (2012) when they perform the nonlinear least square (NLS) estimation. The difference is that we employ the quasi-maximum likelihood estimation, by which we take into account censoring.

$$z^*_{t;i,j,k} = \begin{cases} \dfrac{m - \left(\left(\beta\delta R_j\right)^{\frac{1}{1-\alpha}} - 1\right)\omega}{R + \left(\beta\delta R_j\right)^{\frac{1}{1-\alpha}}} + \epsilon_{i,j,k}, & \text{if } t = 0 \\ \dfrac{m - \left(\left(\delta R_j\right)^{\frac{1}{1-\alpha}} - 1\right)\omega}{R + \left(\delta R_j\right)^{\frac{1}{1-\alpha}}} + \epsilon_{i,j,k}, & \text{if } t > 0 \end{cases}, \epsilon_{i,j,k} \sim N\left(0, \frac{\sigma_k}{100} m\right).$$

Then the choices are censored at the two corners so that the observed choices are

$$z_{t;i,j,k} = \begin{cases} 0, & \text{if } z^*_{t;i,j,k} \le 0 \\ z^*_{t;i,j,k}, & \text{if } 0 < z^*_{t;i,j,k} < \dfrac{m_k}{R_j} \\ \dfrac{m_k}{R_j}, & \text{if } z^*_{t;i,j,k} \ge \dfrac{m_k}{R_j} \end{cases}.$$

This is a two-limit nonlinear censored model, which can be estimated by the quasi-maximum likelihood method.

2.C.2 Results

Table A2.4 reports the estimates of the parameters from the specifications with magnitude-invariant parameters and with magnitude-specific parameters, respectively.

Table A2.5 presents the estimates of the parameter differences between magnitudes. We find a significant magnitude effect on the discount rate. The magnitude effect on the exponent parameter, α, is only significant between the magnitudes of €20 and €40. This is reasonable since we find a strong magnitude effect on the background consumption and the HARA intercept parameter (i.e. ω), which have explained most of the magnitude effects on intertemporal substitutability.

Table A2.4. Discounting and Curvature Parameter Estimates in the Aggregate-Level Estimation with the HARA Specification

Model:	Nonlinear Censored	Nonlinear Censored				
Magnitude:	All	€20	€40	€60	€80	€160
Present bias: $\hat{\beta}$	1.000	0.994	1.002	1.000	1.000	0.997
	(0.004)	(0.010)	(0.005)	(0.005)	(0.005)	(0.005)
Discount factor over four weeks: $\widehat{\delta^{\tau}}$	0.953	0.952	0.960	0.962	0.965	0.967
	(0.004)	(0.008)	(0.003)	(0.003)	(0.004)	(0.004)
Curvature parameter: $\hat{\alpha}$	0.997	0.964	0.984	0.982	0.984	0.986
	(0.003)	(0.009)	(0.006)	(0.004)	(0.003)	(0.002)
Background consumption and HARA intercept: $\hat{\omega}$	9.556	10.713	15.790	22.551	30.834	67.814
	(2.046)	(2.209)	(3.364)	(4.461)	(5.896)	(12.598)
S.d. of the error term: $\hat{\sigma}$	63.948	78.491	68.110	62.147	62.426	65.890
	(4.843)	(6.835)	(5.468)	(4.771)	(5.010)	(5.700)
Log-likelihood	-13136.00			-12836.01		
Observations	7,105			7,105		
Uncensored	1,969			1,969		
Clusters	203			203		

Notes: Quasi-Maximum Likelihood estimators. Column 1: assuming that parameters are the same across magnitudes. Column 2-6: assuming that parameters vary with magnitudes. Clustered standard errors in parentheses. Standard errors calculated via the delta method.

Table A2.5: Estimates of Parameter Differences between Magnitudes in the HARA Specification

Magnitude:	€40 - €20	€60 - €40	€80 - €60	€160 - €80	€60 - €20	€80 - €40	€160 - €60
Present bias: $\hat{\beta}$	0.00807	-0.00180	0.00038	-0.00303	0.00572	-0.00197	-0.00265
	(0.00552)	(0.00206)	(0.00241)	(0.00249)	(0.00578)	(0.00255)	(0.00296)
Discount factor over four weeks: $\widehat{\delta^{\tau}}$	0.00734	0.00252*	0.00258	0.00267*	0.00986*	0.00510***	0.00526**
	(0.00570)	(0.00138)	(0.00169)	(0.00141)	(0.00574)	(0.00176)	(0.00215)
Curvature parameter: $\hat{\alpha}$	0.02007***	0.00097	0.00192	-0.00219	0.01827**	0.00012	0.00411
	(0.00737)	(0.00370)	(0.00225)	(0.00170)	(0.00726)	(0.00459)	(0.00269)
Background consumption and HARA intercept: $\hat{\omega}$	5.077***	6.761***	8.283***	36.981***	11.838***	15.044***	45.263***
	(1.959)	(1.992)	(2.471)	(7.505)	(2.932)	(3.551)	(8.805)

Notes: Estimates of parameter differences are inferred from the quasi-maximum likelihood estimates by the delta method. Parameters are separately set for each magnitude among €20, €40, €60, €80 and €160. There are 1,421 observations (203 clusters) for each magnitude. Clustered standard errors in parentheses. Standard errors calculated via the delta method. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

As we do for the CRRA specification, we use the estimates to predict choices in the 35 questions for both the Present Group and the Delayed Group. Table A2.6 presents the marginal effects of allowing parameters to vary with the magnitude: in Row 1-2, we allow δ to vary with the magnitude of the decisions (as indicated by the column title), but control other two parameters to be the value estimated from the magnitude of €20; in Row 3-4, we allow α and ω to vary with the magnitude. The results show that the marginal effect of allowing α and ω to vary with the magnitude is comparable with the marginal effect of allowing δ to vary. It is consistent with our finding in the Tobit estimation, which implies that the magnitude effect on intertemporal substitutability is at least equally important as the magnitude effect on the discount rate.

Table A2.6: Marginal Effects of Allowing a Parameter to Vary with Magnitudes in the HARA Specification

	Magnitude:	€40	€60	€80	€160
Parameter values used in prediction:	$\delta_k, \alpha_1, \omega_1$ (Delayed):	13.9	22.0	25.9	28.0
	$\delta_k, \alpha_1, \omega_1$ (Present):	17.3	22.1	26.0	28.1
	$\delta_1, \alpha_k, \omega_k$ (Delayed):	8.5	19.4	26.1	36.5
	$\delta_1, \alpha_k, \omega_k$ (Present):	10.9	21.7	29.0	40.2

Notes: The changes in choices predicted by the HARA nonlinear censored model using the parameter values indicated by the row title compared with $(\beta_1, \delta_1, \alpha_1, \omega_1)$, for the two groups separately. k in the row titles stands for the magnitude in the column title. For instance, the first cell in the first row is the difference between the choices made in the seven decisions with the magnitude of €40 predicted by the model with parameter values $(\beta_1, \delta_k, \alpha_1, \omega_1)$ and those predicted by the model with parameter values $(\beta_1, \delta_1, \alpha_1, \omega_1)$. In other words, it is the marginal effect of allowing δ to vary with the magnitude from €20 to €40.

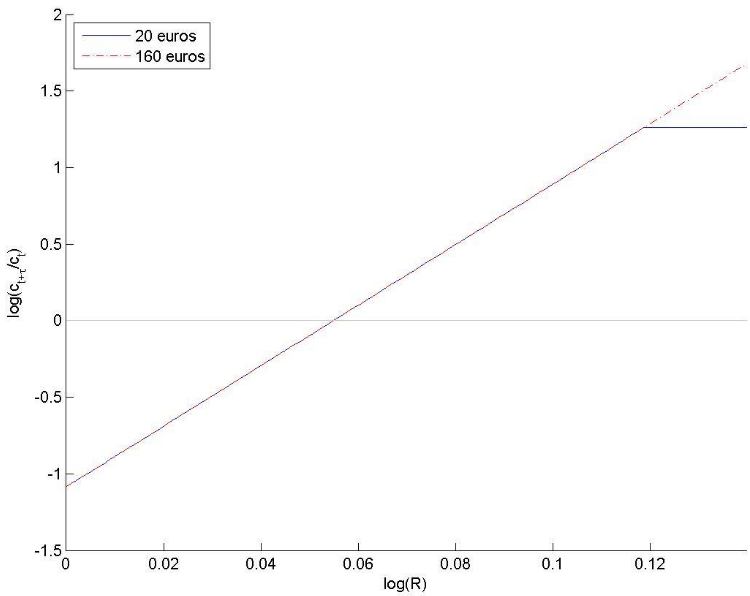
2.D A Simulation of the Mental-Accounting Fudenberg-Levine Model

We perform a simulation according to the mental-accounting Fudenberg-Levine model described in Section 2.5.3. Figure A2.1 shows how the dependent variable in our Tobit estimation, $\ln\left(\frac{z_t+\omega}{z_{t+\tau}+\omega}\right)$, changes with the independent variable, $\ln R$.

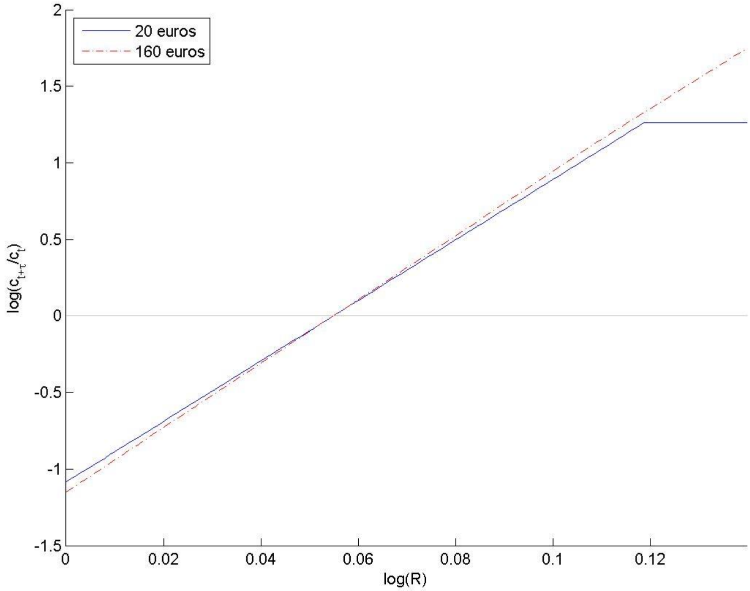
Figure A2.1(a) displays the curves in case the true model is specified by equation (2.1). Note that the slope of the curve (when it is not censored) is $\frac{1}{1-\alpha}$. Thus, a greater slope stands for a larger intertemporal substitutability. The horizontal-axis intercept is $-\ln\delta$. Thus a greater horizontal-axis intercept stands for a smaller discount factor.

Figure A2.1(b) displays the curves in case the true model is the mental-accounting Fudenberg-Levine model. When the budget is €20, both rewards are taken as pocket cash, so the curve is a straight line, as same as predicted by (2.1).

However, when the budget gets larger, the curve is with a greater slope. Therefore, we observe a positive relation between intertemporal substitutability and the size of budget. On the other hand, the horizontal-axis intercept does not change with the stake, suggesting that the model cannot explain magnitude effect on the discount factor.



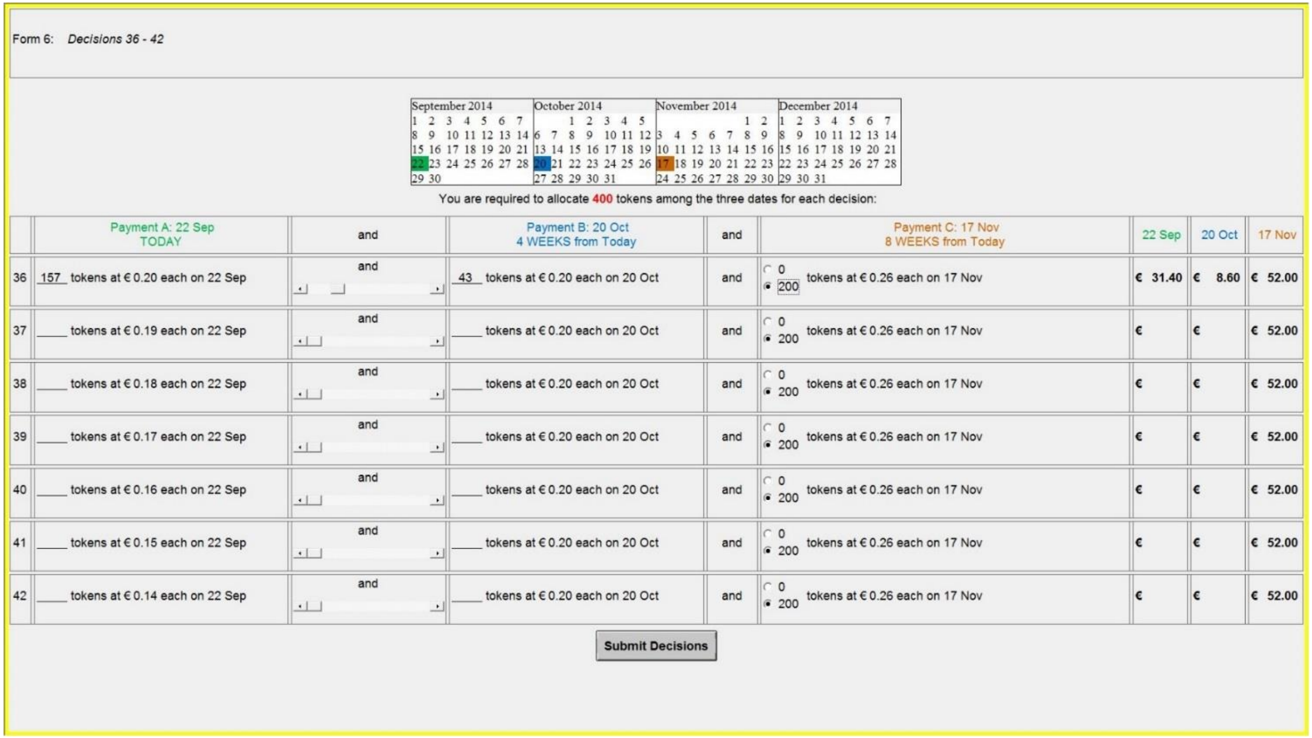
(a) Prediction of the model specified by equation (2.1)



(b) Prediction of the extended Fudenberg-Levine model

Figure A2.1: Simulated Relationships between the Dependent Variable and the Independent Variable of the Tobit Estimation

2.E Decision Forms in Part II



Form 6: *Decisions 36 - 42*

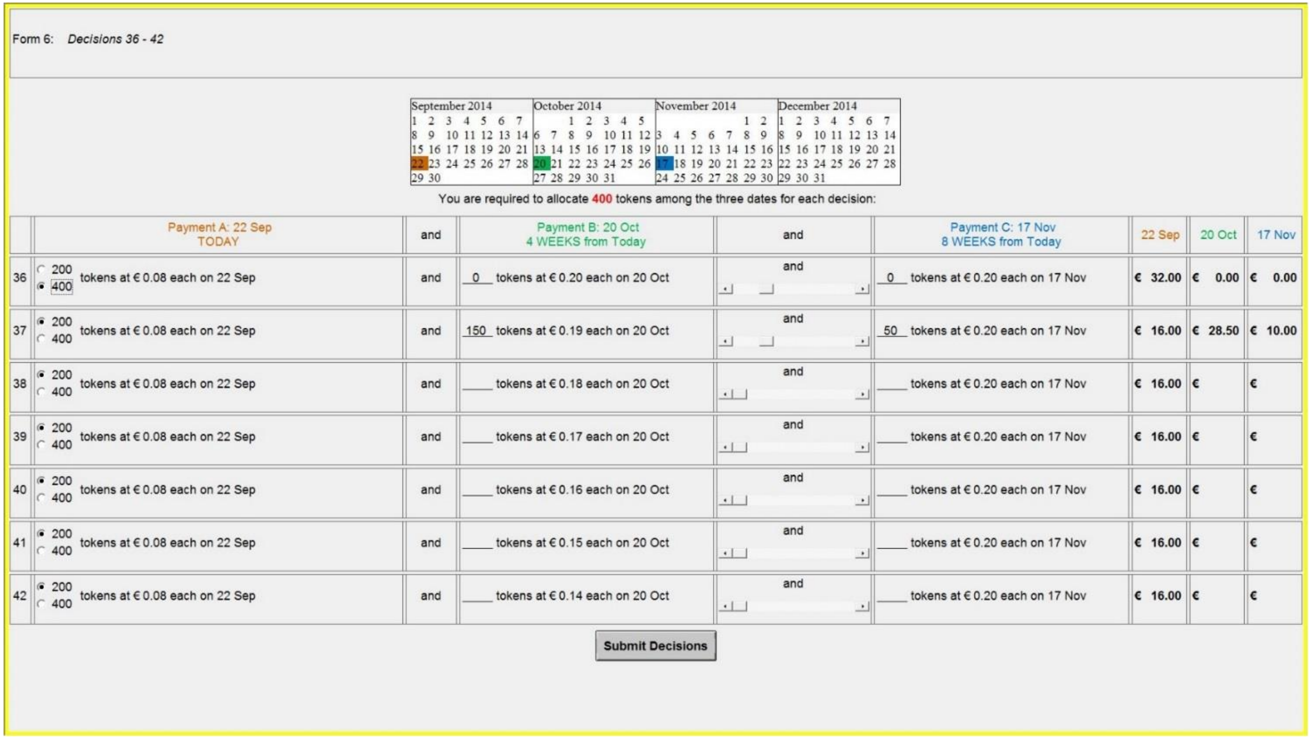
September 2014	October 2014	November 2014	December 2014
1 2 3 4 5 6 7	1 2 3 4 5	1 2	1 2 3 4 5 6 7
8 9 10 11 12 13 14	6 7 8 9 10 11 12	3 4 5 6 7 8 9	8 9 10 11 12 13 14
15 16 17 18 19 20 21	13 14 15 16 17 18 19	10 11 12 13 14 15 16	15 16 17 18 19 20 21
22 23 24 25 26 27 28	20 21 22 23 24 25 26	17 18 19 20 21 22 23	22 23 24 25 26 27 28
29 30	27 28 29 30 31	24 25 26 27 28 29 30	29 30 31

You are required to allocate **400** tokens among the three dates for each decision:

	Payment A: 22 Sep TODAY	and	Payment B: 20 Oct 4 WEEKS from Today	and	Payment C: 17 Nov 8 WEEKS from Today	22 Sep	20 Oct	17 Nov
36	157 tokens at € 0.20 each on 22 Sep	and	43 tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€ 31.40	€ 8.60	€ 52.00
37	____ tokens at € 0.19 each on 22 Sep	and	____ tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
38	____ tokens at € 0.18 each on 22 Sep	and	____ tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
39	____ tokens at € 0.17 each on 22 Sep	and	____ tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
40	____ tokens at € 0.16 each on 22 Sep	and	____ tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
41	____ tokens at € 0.15 each on 22 Sep	and	____ tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
42	____ tokens at € 0.14 each on 22 Sep	and	____ tokens at € 0.20 each on 20 Oct	and	0 / 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00

Submit Decisions

(a) The Present Group



(b) The Delayed Group

Figure A2.2: Interface of a Typical Decision Form in Part II

3 Measuring Preferences over Intertemporal Profiles: Magnitude Effect and All-Sooner Effect

Many economic decisions require people to choose between intertemporal profiles. An *intertemporal profile* is a combination of outcomes that occur at different points in time (e.g. 20 euros today *and* 20 euros in a month is a two-outcome profile). Individuals who select pension plans are choosing among profiles of income over their lifetime. College graduates who decide whether to look for a job or to continue their education are choosing among profiles of income and effort.

Despite the importance of understanding choices between intertemporal profiles, existing experimental measurements of such choices are mostly confined to a special case: choices between *single dated rewards* (i.e. profiles with only one non-zero outcome, such as 20 euros in a month, as opposed to *mixed profiles*, which have more than one non-zero outcome). A limitation of measurement methods using single dated rewards only is that, even if a simple discounting model with only two parameters, a discount rate and a utility curvature, is assumed, those parameters cannot be identified simultaneously. One needs to identify the utility curvature by measuring risk preferences. This joint elicitation method has been questioned since recent evidence showed that utility over time is different from utility under risk: the utility curvatures over time and under risk are uncorrelated at the individual level (Andreoni and Sprenger, 2012a), quantitatively different (Abdellaoui et al., 2013; Miao and Zhong, 2015), and change with stake in opposite directions (Sun and Potters, 2016).

In this paper, we introduce a new method for measuring intertemporal preferences. The idea is simple: one just directly measures preferences over intertemporal profiles. When a *classic discounting model* is assumed (i.e. the intertemporal utility function is composed of a stationary period utility function and a magnitude-independent discount function), the method measures the discount rate and the utility curvature simultaneously. Compared to previous methods (Abdellaoui et al. 2010, Andreoni and Sprenger 2012a, and Cheung 2015a), our method has the advantage of being parameter-free, incentive compatible and independent of time horizon effects at the same time. Most importantly, the method requires weak assumptions on preferences to be measured. Hence, it allows studies of a wider range of models, including those having non-stationary period utility functions (Benhabib et al., 2010), those having a magnitude-dependent discount function (Noor, 2011), and those having an additively non-separable utility function (Holden and Quiggin, 2017).

We then apply the method to investigate how people make choices between two-outcome intertemporal profiles of monetary income. A two-outcome intertemporal profile of monetary income consists of a reward on a sooner date (*sooner reward*) and a reward on a later date (*later reward*). Existing models that provide similar predictions on choices between single dated rewards have different predictions on choices between two-outcome profiles. We focus on two attributes of preferences which distinguish models from each other: how choices between two-outcome profiles change with the scale of outcomes (*the magnitude effect*), and whether people are overly impatient in case it is possible to have all rewards on a sooner date (*the all-sooner effect*).

The study on how choices between intertemporal profiles change as all outcomes are scaled up, is an extension of the well-known magnitude effect over single dated rewards. Studies using single dated rewards found that a *pure later reward* (i.e. a two-outcome profile with zero outcome on the sooner date and positive outcome on the later date) that is equally good as a *pure sooner reward* (i.e. a two-outcome profile with positive outcome on the sooner date and zero outcome on the later date) is more favored when the two rewards are increased by the same ratio (e.g. Thaler, 1981; see Andersen et al., 2013 for a review).[19] But little is known about how choices between general profiles change. Models aimed at explaining the magnitude effect over single dated rewards provide qualitatively different predictions on choices between intertemporal profiles (Loewenstein and Prelec, 1992; Noor, 2011; Baucells and Heukamp, 2012; Benhabib et al. 2010), and hence it is interesting to know which model is more consistent with observation. In specific, those models differ in predictions on two channels of the magnitude effect: the discount rate and the utility curvature. The two channels represent patience and perceived fungibility, respectively, and can predict opposite directions of change as outcomes are scaled up.

The second question, whether people are overly impatient when a pure sooner reward is available, concerns a channel of impatience other than discounting. Let (x, y) stands for the profile of receiving €x on a sooner date (e.g. tomorrow) and €y on a later date (e.g. in five months). A decision-maker who is indifferent between $(28,10)$ and $(20,20)$ reveals that she is impatient; she is willing to exchange €10 on the later date for €8 on the sooner date. The question here is whether she prefers $(36,0)$ to $(20,20)$. If so, she is willing to exchange later rewards for sooner rewards at a higher price, suggesting that she is more impatient when one of the options is a pure sooner reward. This phenomenon is predicted by a model with a fixed cost of waiting (introduced by Benhabib et al. 2010 and extended to two-outcome profiles), because a fixed cost incurred whenever one needs to wait implies a preference for pure sooner rewards over mixed profiles. The phenomenon in question is analogous to the certainty effect (Allais, 1953; Tversky and Kahneman, 1981), which is about a preference for certain outcomes over risky prospects. It has implications for estimating preferences in scientific studies and prescriptive consultation: omitting the all-sooner effect leads to an overestimation of the discount rate and an underestimation of the utility curvature.

[19] For instance, in Halevy (2015), a median subject is indifferent between \$10 today and \$10.59 in a week, implying a weekly discount rate of 5.9%, but she is indifferent between \$100 today and \$103.68 in a week, implying a weekly discount rate of 3.7%.

Several experiments also involved choices between intertemporal profiles but addressed different questions than the current paper. Andreoni and Sprenger (2012a) and Cheung (2015a) found evidence on convexity of preferences. Abdellaoui et al. (2010) examined the gain-loss asymmetry. Attema et al. (2016) introduced a measurement method of the discount function without measuring the period utility function. Andreoni and Sprenger (2012b), Miao and Zhong (2015), and Cheung (2015b) investigated the interaction of time and risk.[20]

This paper makes four contributions to the literature on intertemporal choices. First, the paper introduces a new method for measuring preferences over intertemporal profiles. The method does not require parametric assumptions of utility functions or discount functions, nor does it require the assumption of a stationary period utility function and a magnitude-independent discount function. It even allows additive non-separability. Therefore, the method provides an option to measure preferences under minimum assumptions. This implies that a wider range of models can be studied, and that misspecifications are less likely to occur.

Second, the paper is the first to investigate whether people have a preference for pure sooner rewards over mixed profiles, i.e. whether the all-sooner effect exists. The all-sooner effect suggests a different channel of impatience than discounting. It also has implications for the measurement of preferences: if one measures preferences with single dated rewards, the existence of the all-sooner effect and the ignorance of it imply an overestimation of the discount rate and an underestimation of the utility curvature.

Third, the paper examines the channels of the magnitude effect non-parametrically. The different channels of the magnitude effect determine in what directions choices change with the stake. Sun and Potters (2016) performed the first investigation into the magnitude effect on choices between intertemporal

[20] There are also studies using field data to investigate properties of preferences over intertemporal profiles (for instance, Laibson et al., 2007, estimate the quasi-hyperbolic discount function).

profiles, and disentangled channels by assuming a parametric model. In contrast, the current paper explores the channels of the magnitude effect without assuming any specific functional form.

Lastly, the paper provides a simple model that captures all characteristics found in our experiment, and it is flexible enough to represent different types of subjects. The model facilitates parametric estimation of preferences in future applications.

We find evidence that the elasticity of intertemporal substitution is larger for higher stakes, but no evidence shows that the generalized discount factor is also larger for higher stakes. This suggests that scales of outcomes affect choices mainly through the perceived fungibility of rewards across time, rather than the patience in general. We find evidence of the all-sooner effect, suggesting that people have a preference for pure sooner rewards over mixed profiles.

Our results provide implications for preference estimation with single dated rewards or small incentives. If the all-sooner effect is omitted, estimating preferences with single dated rewards will lead to an overestimation of the discount rate and an underestimation of the utility curvature. If the magnitude effect is omitted, estimating preferences with small stakes will lead to an overestimation of the utility curvature.

The remaining part of the paper is structured as follows: Section 3.1 introduces some concepts and notations. Section 3.2 presents the measurement method. Section 3.3 describes the theoretical framework of our application. In particular, the predictions of existing models regarding the magnitude effect and the all-sooner effect are provided. The experimental design is presented in Section 3.4. Section 3.5 shows the results. A tractable and flexible model which accommodates all characteristics found in our results is also provided. A discussion follows in Section 3.6.

3.1 Background

We consider two-outcome intertemporal profiles, each of which is represented by $(x, y) \in \mathbb{R}_2^+ \equiv [0, +\infty)^2$, where x denotes the outcome at time point t_1 (i.e. the *sooner reward*), y denotes the outcome at time point $t_2 = t_1 + \tau$ (i.e. the *later reward*), and τ is the delay. We fix t_1 and t_2 to control the time horizon effects in this study, though our measurement method can be used in situations with different front-end delays and delays. [21]

We examine preferences $\succcurlyeq$ over intertemporal profiles in $\mathbb{R}_2^+$. $\sim$ denotes indifference as usual. We restrict our attention to preferences that are represented by a utility function, $U(x, y)$, which is continuous, strictly increasing and differentiable on the *sooner boundary* $[0, +\infty) \times \{0\}$ and in the remaining part of the space $[0, +\infty) \times (0, +\infty)$, respectively. By allowing the preference to be discontinuous or non-monotonic between the sooner boundary and the interior, we accommodate the fixed-cost-of-waiting model by Benhabib et al. (2010).[22]

Following the convention, the *marginal rate* of *substitution* of the sooner reward for the later reward (henceforth MRS) at profile (x, y) is

(3.1) $$MRS(x, y) = \frac{U_2'(x, y)}{U_1'(x, y)}.$$

The MRS at profile (x, y) measures how much sooner reward a decision-maker is willing to exchange for one unit of later reward, which reflects the decision-maker's patience locally at the profile. She is more patient locally if she has a larger MRS.

Suppose (x_0, y_0) is a profile in $\mathbb{R}_2^+$, and L is a ray in $\mathbb{R}_2^+$ from the origin with the income ratio $\frac{x}{y} = \lambda \in [0, +\infty]$. By continuity and strict monotonicity,

[21] There are numerous studies on time horizon effects, including the discussions on the present bias, time inconsistency and sub-additivity in discounting. See, for instance, a recent synthesis by Dohmen et al. (2017). In Section 3.6, we discuss how time horizon effects can be combined with our findings.

[22] In Section 3.6, we discuss another set of assumptions where monotonicity holds for the entire space but transitivity can be violated. In that case, there are also models consistent with our empirical results.

there must be a profile $(x_1, y_1) \in L$ such that $(x_0, y_0) \sim (x_1, y_1)$. The *average rate of substitution* of the sooner reward for the later reward (henceforth ARS) between the profile (x, y) and the ray L is

(3.2) $$ARS(x_0, y_0, \lambda) = -\frac{x_1 - x_0}{y_1 - y_0}.$$

The ARS reflects the patience revealed from choices between (x_0, y_0) and profiles on L. The decision-maker is more patient if she has a larger ARS.

3.2 Measurement Method

We introduce a new method of measuring preferences over intertemporal profiles in a parameter-free way. The method does not require the assumption of a stationary utility function and a magnitude-independent discount function. It even allows the preference to be additively non-separable.

The main idea of the method is to measure indifference curves through a set of pre-determined profiles. For measuring each indifference curve, a few indifference relations are elicited between one pre-determined profile and profiles with a few pre-determined income ratios. Figure 3.1 illustrates how an indifference curve is measured. The indifference curve through the predetermined profile (x_0, y_0) is measured by eliciting the indifference relations between (x_0, y_0) and each of (x_1, y_1), (x_2, y_2), … , (x_J, y_J), where $\lambda_j = \frac{x_j}{y_j}, j = 1, \dots, J$ are pre-determined by the experimenter and satisfy $0 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_J \leq \infty$. By connecting all those equally good profiles, an indifference curve is obtained.

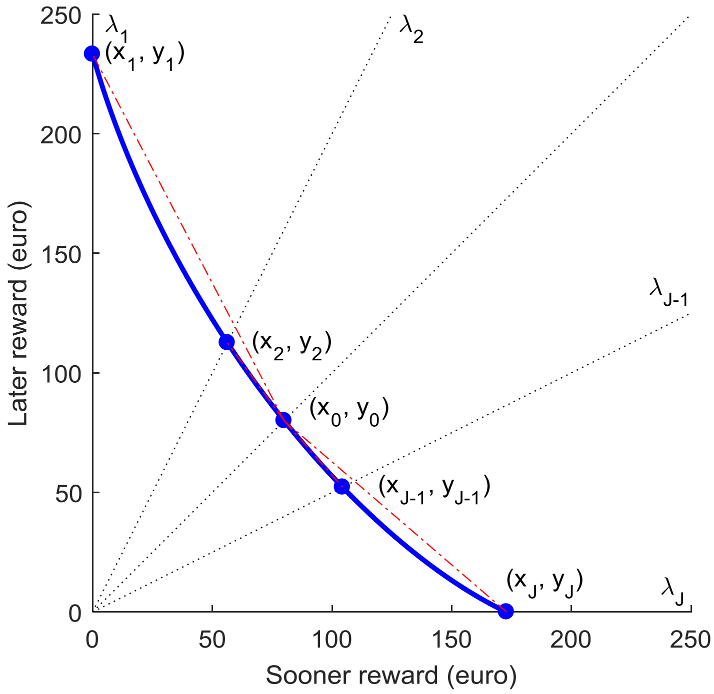


Figure 3.1: Parameter-Free Method to Measure Preferences over Intertemporal Profiles

To elicit each of the indifference relation, a choice list is used. In one choice list, M binary choice problems are presented. One of the options in all those problems is the balanced profile, (x_0, y_0). The other option moves along the line towards the origin. For instance, in the choice list eliciting the indifference relation between (x_0, y_0) and (x_j, y_j), one of the options is always (x_0, y_0), while the other option changes from (x_j^1, y_j^1) to (x_j^M, y_j^M), where $\frac{x_j^1}{y_j^1} = \cdots = \frac{x_j^M}{y_j^M} = \lambda_j$ and either $x_j^1 > x_j^2 > \cdots > x_j^M$ or $y_j^1 > y_j^2 > \cdots > y_j^M$ holds. Since the varying option is getting worse as a subject moves down the list, the indifference relation can be identified by looking at the switch point. Figure 3.2 shows an example of a choice list. The RIGHT option is always (€20, €20), while the LEFT option changes from (€0, €60) to (€0, €20), all of which have an income ratio of zero.

The measurement method has some advantages compared to previous methods. Compared to the chained procedure used by Abdellaoui et al. (2010), our elicitation is incentive compatible, and does not rely on additive separability. Incentivization is useful especially for studies on the magnitude effect, because hypothetical studies reported a much larger size of magnitude effect than real stake ones. The flexibility of our method enables the study of a wider range of models, including those with non-stationary or additively non-separable utility functions. Compared to the Convex Time Budget method (Andreoni and Sprenger 2012a) which measures the choice correspondence, our method directly measures the preference relations between pairs of profiles, and hence does not rely on parametric assumptions of utility functions. Compared to the choice list used by Cheung (2015a) which varies the delay, our method fixes the front-end delay and the delay, so that the measurement is independent of time horizon effects. This implies researchers do not need to assume exponential discounting or any other discount function to obtain estimates of discount rates. Compared to Attema et al. (2016) which measures discount functions without measuring utility functions, our method estimates the discount function and the utility function simultaneously, and again it does not require the assumption of stationary period utility functions.

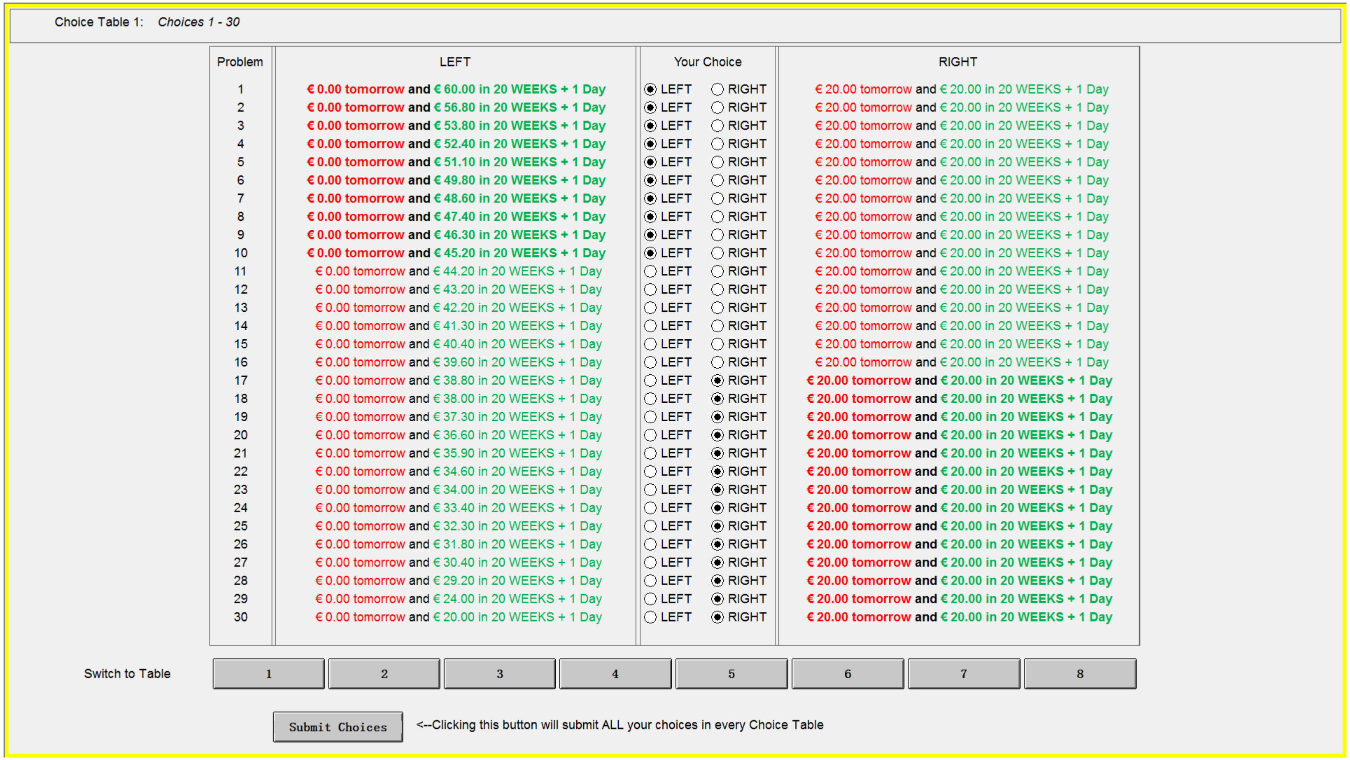
Choice Table 1: *Choices 1 - 30*

Problem	LEFT	Your Choice	RIGHT
1	€ 0.00 tomorrow and € 60.00 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
2	€ 0.00 tomorrow and € 56.80 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
3	€ 0.00 tomorrow and € 53.80 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
4	€ 0.00 tomorrow and € 52.40 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
5	€ 0.00 tomorrow and € 51.10 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
6	€ 0.00 tomorrow and € 49.80 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
7	€ 0.00 tomorrow and € 48.60 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
8	€ 0.00 tomorrow and € 47.40 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
9	€ 0.00 tomorrow and € 46.30 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
10	€ 0.00 tomorrow and € 45.20 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
11	€ 0.00 tomorrow and € 44.20 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
12	€ 0.00 tomorrow and € 43.20 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
13	€ 0.00 tomorrow and € 42.20 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
14	€ 0.00 tomorrow and € 41.30 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
15	€ 0.00 tomorrow and € 40.40 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
16	€ 0.00 tomorrow and € 39.60 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
17	€ 0.00 tomorrow and € 38.80 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
18	€ 0.00 tomorrow and € 38.00 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
19	€ 0.00 tomorrow and € 37.30 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
20	€ 0.00 tomorrow and € 36.60 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
21	€ 0.00 tomorrow and € 35.90 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
22	€ 0.00 tomorrow and € 34.60 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
23	€ 0.00 tomorrow and € 34.00 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
24	€ 0.00 tomorrow and € 33.40 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
25	€ 0.00 tomorrow and € 32.30 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
26	€ 0.00 tomorrow and € 31.80 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
27	€ 0.00 tomorrow and € 30.40 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
28	€ 0.00 tomorrow and € 29.20 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
29	€ 0.00 tomorrow and € 24.00 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day
30	€ 0.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day	◯ LEFT ◯ RIGHT	€ 20.00 tomorrow and € 20.00 in 20 WEEKS + 1 Day

Switch to Table 1 2 3 4 5 6 7 8

Submit Choices <--Clicking this button will submit ALL your choices in every Choice Table

(a) Before any option is chosen



(b) After LEFT is chosen for Choice Problem 10 and RIGHT is chosen for Choice Problem 17

Figure 3.2: The Interface Before and After Some Options are Chosen

One disadvantage of our method is that it requires a large number of choices for identifying an indifference curve. This is caused by two reasons. First, for identifying each indifference relation, a choice list with tens of binary choice problems is needed. However, when ordered choice lists are used and computerized programs force subjects to switch at most once in each list, subjects only need to make one actual decision for each list (see Section 3.4.1). This could substantially reduce the workload of subjects. Second, measuring an indifference curve with $(J + 1)$ points requires J choice lists. This is a price for non-parametric measurement and can be reduced if some parametric assumptions are imposed.

3.3 Measuring Preferences and Testing Models: Theory

We apply our method to measure preferences over two-outcome intertemporal profiles. Since our method requires weak assumptions on the model to be measured, the measurement is especially suitable for testing models of intertemporal choices, including those beyond a classic discounting model.

We are interested in testing models of intertemporal choices because most existing models provide similar predictions to choices between single dated rewards but qualitatively different predictions to choices between intertemporal profiles.

We focus on two attributes of preferences: how choices change as outcomes are scaled up (magnitude effect) and whether people are overly impatient when a pure sooner reward is available (all-sooner effect). Those two attributes are key to predicting choices, have clear economic meanings and apply to all models satisfying weak assumptions. Importantly, existing models have different predictions on the two attributes, so they can be used to distinguish models.

In the following subsections, we define and interpret the magnitude effect and the all-sooner effect, and then provide predictions of existing models on the two attributes.

3.3.1 Channels of the Magnitude Effect

In order to study how choices over intertemporal profiles change with the stake, we look at two properties of intertemporal preferences.

First, the *generalized discount factor* (henceforth GDF) at the magnitude of x is the MRS at the balanced profile (x, x). Formally,

(3.3) $$\delta(x) \equiv MRS(x, x) = \frac{U_2'(x, x)}{U_1'(x, x)}.$$

If the preference can be represented by a classic discounting model, i.e. it is representable by an additively separable utility function with a stationary utility function and a magnitude-independent discount function, as

(3.4) $$U(x, y) = u(x) + D(\tau) \cdot u(y),$$

where D is the discount function, then it holds that

(3.5) $$\frac{U_2'(x, x)}{U_1'(x, x)} = D(\tau).$$

Thus, the GDF reduces to the regular discount factor of a delay τ in a classic discounting model.

Second, the *elasticity of intertemporal substitution* (henceforth EIS) at a profile (x, y) is the relative change in income ratio as a response to a relative change in the MRS given the utility level at (x, y). Formally,

(3.6) $$\varepsilon(x, y) \equiv \frac{1}{\frac{\partial}{\partial Q} \ln \frac{U_2'\big(x(Q, u), y(Q, u)\big)}{U_1'\big(x(Q, u), y(Q, u)\big)}},$$

where $Q \equiv \ln\left(\frac{x}{y}\right)$ and $u = U(x, y)$.

Here, Q is the income ratio, and $\ln \frac{U_1'}{U_2'}$ is the MRS, which is equal to the interest rate when utility is optimized. As usual, the elasticity measures the sensitivity of quantity ratios to relative price. The larger the EIS, the more sensitive choices are to changes in interest rate.

Graphically, the EIS is a measure of the local curvature of the indifference curve: the larger the elasticity, the less convex the indifference curve is around (x, y). In other words, it measures the rate of change in the MRS as the profile moves along the indifference curve.

If the preference is represented by a classical discounting model with a power utility function, i.e.

(3.7) $$U(x,y) = \frac{x^\alpha}{\alpha} + D(\tau) \cdot \frac{y^\alpha}{\alpha},$$

then it holds

(3.8) $$\varepsilon(x,y) = \frac{1}{1-\alpha}.$$

Hence, the EIS reduces to a positive transformation of the curvature parameter α. The linear utility model corresponds to $\alpha = 1$, or $\varepsilon = \infty$, and the Leontief utility model corresponds to $\alpha = \infty$, or $\varepsilon = 0$.

We focus on the two properties of preferences for four reasons. First of all, those properties have important economic meanings. As long as the preference is convex, the GDF corresponds to the interest rate at which the decision maker is willing to hold a balanced profile. In this sense, it reflects the decision maker's average patience at the corresponding magnitude. It also predicts how people make choices between close-to-balanced profiles. On the other hand, the EIS measures the perceived fungibility between the sooner reward and the later reward. It also predicts how choices between unbalanced profiles differ from those between close-to-balanced profiles.

Besides their economic relevance, the two properties also have the advantage of generality. They can be obtained for any preference under our weak assumptions. They do not rely on the assumption of a specific functional form, nor on the existence of a stationary period utility function or a magnitude-independent discount function. They do not even require additive separability. Hence, empirical results based on the two properties are robust, and tests based on them can be applied to a wide range of models.

Though the definitions of the GDF and the EIS are independent of model assumptions, they reduce to familiar parameters when some popular assumptions are imposed. Thereby, with those assumptions, findings about the two properties are comparable to those in previous studies.

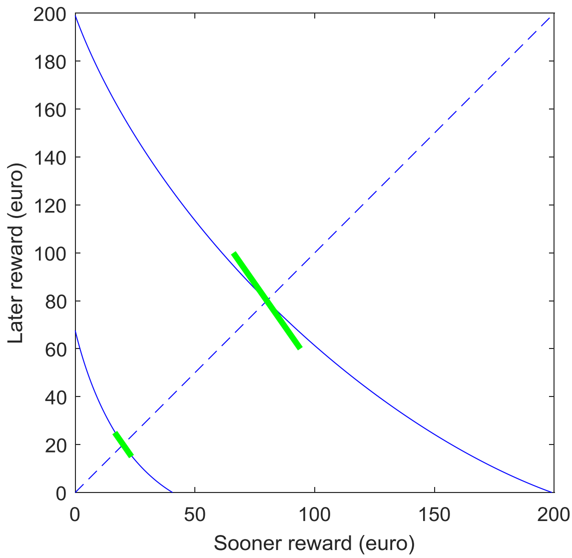
Most importantly, how the GDF and the EIS change with the scale of outcomes have different implications for how choices change with the scale. Figure 3.3 shows how choices change as both profiles are scaled up, if the GDF or the EIS is larger for a higher stake, respectively. In the two panels, each line segment represents a binary choice problem, and the two endpoints of each line stand for the two options. If the GDF is larger for a higher stake, as shown in Figure 3.3(a), the decision-maker is indifferent between the two profiles indicated by the line segment when their stakes are low, but prefers the profile with a larger proportion of later reward when all outcomes are scaled up. This suggests that she appears to be more patient in choices between close-to-balanced profiles when stakes are higher. If the EIS is larger when the profile is scaled up, as shown in Figure 3.3(b), when outcomes are scaled up, the decision-maker care less about smoothing but more about the discounted sum of amounts, and hence the decision-maker appears to be more patient in choices between *future-leaning profiles* (i.e. two-outcome profiles of which the later reward has a greater amount than the sooner reward) and less patient in choices between *present-leaning profiles* (i.e. two-outcome profiles of which the sooner reward has a greater amount than the later reward). The two channels can work simultaneously, making the decision-maker more patient in choices between future-leaning profiles, but how choices between present-leaning profiles change depends on the net effect of the two channels.

3.3.2 *The All-Sooner Effect*

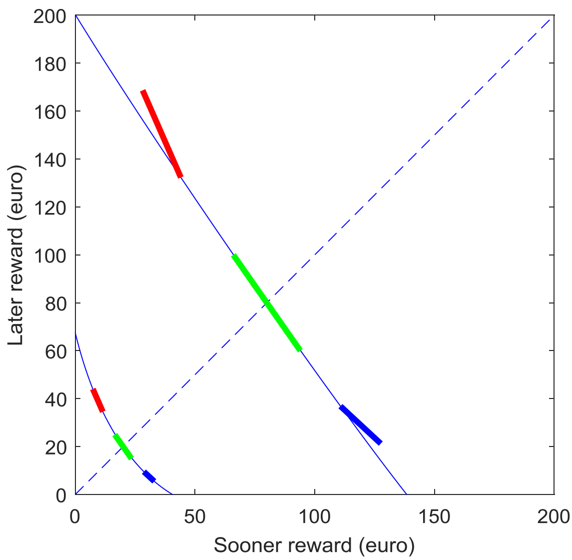
The *all-sooner effect* exists if there exist a mixed profile, (x_1, y_1), a pure sooner reward, $(x_0, 0)$, and a parameter $\lambda \in (0,1)$, such that $(x_1, y_1) \succcurlyeq (\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$ and $(x_0, 0) \succ (\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$.

The intuition of the all-sooner effect is that people have a preference for pure sooner rewards. A pure sooner reward brings a higher utility than mixed profiles close by. One consequence is that people appear to be less patient when one of the options is a pure sooner reward. Figure 3.4 shows an example: (x_1, y_1), $(\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$ and $(x_0, 0)$ are three profiles on a straight line, (x_1, y_1), $(\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$ and $(x_2, 0)$ are equally good, and $x_2 < x_0$. By monotonicity, $(x_0, 0)$ is more favored than the other three profiles, and hence the all-sooner effect exists. As a result, the indifference relation between (x_1, y_1) and $(x_2, 0)$ reveals an ARS of $\frac{x_2 - x_1}{y_1}$, which is less than the ARS revealed by the indifference relation between (x_1, y_1) and $(\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$, or $\frac{x_0 - x_1}{y_1}$. Hence, given the same rate of return, the decision-maker appears to be less patient in a choice problem where a pure sooner reward is available.

The all-sooner effect is analogous to the certainty effect. The certainty effect is about the larger disutility caused by a reduction in probability from a certain prospect than by a reduction in probability from a risky prospect. In contrast, the all-sooner effect is about the larger disutility caused by the delay of some money from a pure sooner reward than by a delay of some money from a mixed profile.



(a) When the GDF becomes larger, people appear to be more patient in choices between close-to-balanced profiles



(b) When the EIS becomes larger, revealed patience changes in different directions for choices in different domains

Figure 3.3: Implications of Different Channels of the Magnitude Effect

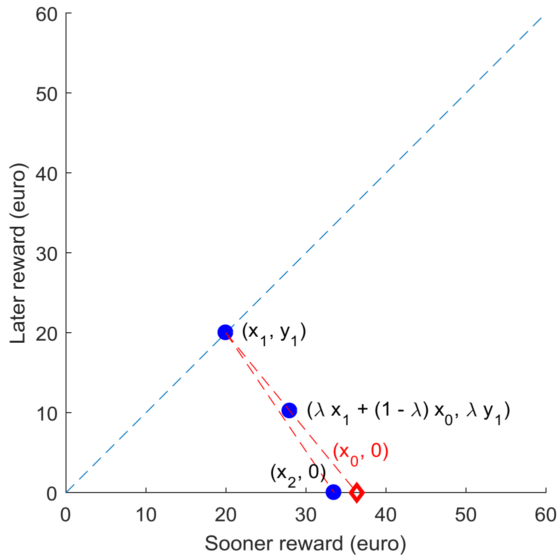


Figure 3.4: The All-Sooner Effect Implies a Violation of Convexity

The all-sooner effect is consistent with the fixed-cost-of-waiting model proposed by Benhabib et al. (2010). They argue that subjects behave as if they incur a fixed cost (of $4, as calibrated in their experiment) when a sooner reward is delayed entirely to the later date, regardless of the amount being delayed. They used the model to explain decreasing impatience and the magnitude effect over single dated rewards. It can also be used to account for the all-sooner effect if the model is extended to intertemporal profiles in the following way: a fixed cost is incurred unless one receives all the money on the earliest possible day. Figure 3.5 shows how the fixed-cost-of-waiting model generates the all-sooner effect. The indifference curve is discontinuous at the sooner boundary; it jumps from $(x_2', 0)$ to $(x_2, 0)$, so that (x_1, y_1), $(\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$ and $(x_2, 0)$ are equally good, while $(x_0, 0)$ is more favored. Thereby, the all-sooner effect exists.

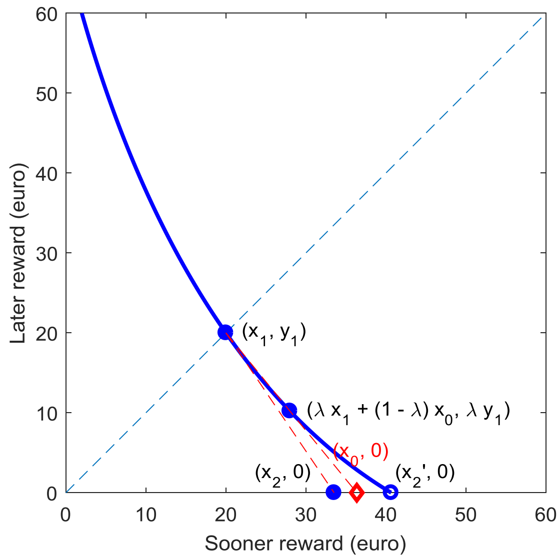


Figure 3.5: A Fixed Cost of Waiting Leads to the All-Sooner Effect

The all-sooner effect has implications for both theory and empirical work. First, it suggests a channel of impatience other than discounting. Impatience between single dated rewards is usually modelled as discounting. While discounting represents a proportional cost of delaying rewards, the all-sooner effect stands for a disproportional cost, including the case of a fixed cost. It is interesting to know whether the impatience between single dated rewards is caused by discounting or the all-sooner effect or both, because the two channels have different implications for choices between intertemporal profiles.

Second, the all-sooner effect implies a violation of convexity. Convexity is assumed by most models and applications on intertemporal choices. Intuitively, convexity says people have a preference for outcome smoothing, i.e. they prefer more balanced profiles. The all-sooner effect violates convexity, because the preference for pure sooner rewards outweighs the preference for smoothing when the difference in income ratio is small. Figure 3.4 shows how the all-sooner effect implies a violation of convexity. By definition, $(x_1, y_1) \succcurlyeq (\lambda x_1 + (1 -$

$\lambda)x_0, \lambda y_1)$ and $(x_0, 0) \succ (\lambda x_1 + (1-\lambda)x_0, \lambda y_1)$. Therefore, the indifference curve cannot be convex between (x_1, y_1) and $(x_2, 0)$.[23]

The all-sooner effect is not ruled out by existing evidence of convexity. In previous studies, convexity was either tested with choices between mixed profiles (Cheung, 2015a) or tested parametrically with the assumption of a uniform curvature (Andreoni and Sprenger, 2012a). Hence, their results are not conflicting with non-convexity near the sooner boundary. Abdellaoui et al. (2010) also found evidence of convexity. We notice that in their results, the median subject is less patient between a pure sooner reward and a mixed profile than between two mixed profiles with the lowest amounts. This is actually in line with the all-sooner effect.[24]

Third, the all-sooner effect also has implications for estimating preferences. Estimating preferences with single dated rewards leads to an overestimation of the discount rate if the all-sooner effect is omitted, because part of the "discounting" is actually due to the all-sooner effect. Estimating preferences with both single dated rewards and some mixed profile leads to an underestimation of the utility curvature if the all-sooner effect is omitted, because the all-sooner effect cancels out part of the convexity.

3.3.3 Predictions of Various Models

In order to investigate what kinds of models provide predictions more consistent with empirical patterns on the magnitude effect and the all-sooner effect, we list out the predictions of various models in terms of the magnitude effect on the GDF and the EIS and the all-sooner effect.

[23] The all-sooner effect is related to outcomes, and not to timing. Thereby, it cannot be explained by any discount function.

[24] They did not perform a formal test on the non-convexity. Besides, their measurement relies on the assumption of additive separability, and entails risks in the sooner rewards. Therefore, their dataset is not ideal for performing a test on the all-sooner effect.

Table 3.1 summarizes the predictions of eight models on the two attributes of our interest. Proofs are provided in Appendix 3.A.

The first three models in Table 3.1 are the most popular ones in the empirical literature. All of them are classic discounting models. The first model assumes a power utility function, the second a power utility function with a background consumption (also called Stone-Geary utility function), and the third an exponential utility function. Since they are all discounting models with a magnitude-independent discount function, they predict a constant GDF. The first model predicts a constant EIS, and the other two predict an EIS decreasing in the stake.

Loewenstein and Prelec (1992) assumed a sub-proportional utility function to account for the magnitude effect over single dated rewards. The model predicts a constant GDF, but the relation between the EIS and the stake is ambiguous. Two examples are given to show that the EIS can be either increasing or decreasing in the stake. If the period utility function is a power function with a positive background consumption, the EIS is decreasing in the stake. If $u(x) = x + \frac{x^\alpha}{\alpha}$, the EIS is increasing in the stake.

Fudenberg and Levine (2006) proposed a bank-nightclub model based on their costly self-control model. The main idea is that people who receive a windfall will balance the impulse of consuming and the long-run benefits of savings if the windfall is not too small, but will spend all the money if the windfall is small. If we take this model as an expression of how people treat windfalls, assume that people do not adjust their consumption plan until they receive the windfall, and omit the effect of the wealth accumulation in the first period on the background consumption in the second period, then we have a classic discounting model with the following period utility function:

$$u(z) = \begin{cases} \dfrac{(z+\omega)^{\alpha}}{\alpha}, & \text{if } z \le \left[(1+\gamma)^{\frac{1}{1-\alpha}} - 1\right]\omega \\ \left(\dfrac{\delta}{1-\delta} + (1+\gamma)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \dfrac{\left(z + \dfrac{\omega}{1-\delta}\right)^{\alpha}}{\alpha} - \gamma\dfrac{(z+\omega)^{\alpha}}{\alpha} - \dfrac{\omega^{\alpha}}{(1-\delta)\alpha}, & \text{if } z > \left[(1+\gamma)^{\frac{1}{1-\alpha}} - 1\right]\omega \end{cases} \tag{3.9}$$

where $\gamma \geq 0$ is the coefficient of the self-control cost, $\delta < 1$ is the discount factor, $\alpha < 1$ is the utility curvature and $\omega > 0$ is the background consumption. When the outcome is small, the period utility function is exactly a power utility function with a small background consumption, ω. When the outcome is larger than a cutoff, the utility function gradually approximates a power utility function with a large background consumption, $\frac{\omega}{1-\delta}$, which is actually the life-long wealth. The model also predicts a constant GDF, but it predicts an upward jump in the EIS.

Unlike the previous two models which assume that changes in stakes only affect the period utility function, Noor (2011) proposed a model with a magnitude-dependent discount function. The idea is that people are extremely impatient toward a very small outcome, but as the size of the outcome increases, they become more patient. The parametric model provided by the author predicts that the GDF increases from 0 to 1. Meanwhile the indifference curve changes from infinitely concave (i.e. only the sooner reward matters) to mildly convex (the same as the discounting model with a power utility function).[25]

A third explanation to the magnitude effect is given by Benhabib et al. (2010) in their fixed-cost-of-waiting model. If a fixed cost is incurred whenever the decision-maker delays some money from the sooner date to the later date, we get a discounting model which is discontinuous on the sooner boundary. The model predicts a constant GDF and a constant EIS, as well as the all-sooner effect.

Lastly, we consider Holden and Quiggin's (2017) zooming model. In their original model, the background consumption is a function of the later reward in

[25] The general form of the model has too few restrictions to provide a definitive prediction about the GDF and the EIS. Baucells and Heukamp's (2012) Probability and Time Tradeoff model is of a similar case. We thereby only consider the parametric form provided by Noor (2011).

a binary choice problem. We adapt the model to choices over intertemporal profiles by modifying the background consumption into a function of the total amount of the rewards in a profile. The model is additively non-separable. Due to the generality of the GDF and the EIS, we are still able to examine the channels of the magnitude effect generated by this model. The model also predicts a decreasing EIS, and meanwhile it predicts a GDF decreasing in the stake.

To sum up, the models we consider differ in their predictions about whether the GDF and the EIS are larger, the same or smaller for higher stakes, and whether the all-sooner effect exists. In order to tell what kinds of models are consistent with empirical patterns, we measure the preferences and focus on the following three empirical questions:

Question 1: Is the EIS the same for different stakes?

Question 2: Is the GDF the same for different stakes?

Question 3: Is the all-sooner effect no-existent?

We measure the ARSs between the balanced profile and one of the four rays from the origin (two interior rays and the two boundaries), both for a lower stake and for a higher stake. With the measures, we test if the GDF and the EIS are larger, the same or smaller for the higher stake than for the lower stake, and if the ARS is smaller near the sooner boundary. By checking those features, we examine what kinds of models better predict observed choices.

Our test of models on intertemporal choices is similar in spirit with Kerschbamer's (2015) test of models on distributional preferences, in the sense that we both test models according to core features of preferences rather than goodness of fit of specific models. The idea is to classify and test models in terms of the characteristics with important economic meanings and high predictive power, keeping the test unaffected by specific assumptions that are mainly for technical convenience.

Table 3.1: Predictions of Various Models in terms of the GDF, the EIS and the All-Sooner Effect

Model	EIS	GDF	All-sooner effect
Power utility function + magnitude-independent discount factor: $$U(x,y)=\frac{x^\alpha}{\alpha}+\delta\frac{y^\alpha}{\alpha}$$ where $\alpha<1$	Constant	Constant	No
Power utility function + positive background consumption + magnitude-independent discount factor: $$U(x,y)=\frac{(x+\omega)^\alpha}{\alpha}+\delta\frac{(y+\omega)^\alpha}{\alpha}$$ where $\alpha<1$ and $\omega>0$	Decreasing in the stake, and converging to $\frac{1}{1-\alpha}$	Constant	No
Exponential utility function + magnitude-independent discount factor: $$U(x,y)=1-\exp(-x)+\delta(1-\exp(-y))$$	Decreasing in the stake, and converging to 0	Constant	No
Loewenstein and Prelec's (1992) sub-proportional utility function: $$U(x,y)=u(x)+\delta u(y)$$ where $$x_1<x_2 \text{ and } k>1 \Rightarrow \frac{u(x_1)}{u(x_2)}\frac{u(kx_2)}{u(kx_1)}>1$$	Depending on $u(\cdot)$. E.g. Decreasing in the stake if $u(x)=\frac{(x+\omega)^\alpha}{\alpha}$, but increasing in the stake if $u(x)=x+\frac{x^\alpha}{\alpha}$	Constant	No
A mental accounting version of Fudenberg and Levine's (2006) bank-nightclub model	First decreasing, then jumping to a larger value, and decreasing again	Constant	No
The parametric form of Noor's (2011) magnitude-dependent discounting model: $$U(x,y)=d^{\frac{t_1}{x^\beta}}\frac{1}{\alpha}x^\alpha+d^{\frac{t_2}{y^\beta}}\frac{1}{\alpha}y^\alpha$$ where $\beta<\alpha<1$ and $d<1$	The indifference curve turns from infinitely concave to mildly convex (with the elasticity of $\frac{1}{1-\alpha}$)	Increasing from 0 to 1	No

Benhabib, Bisin and Schotter's (2010) fixed-cost-of-waiting model: $U(x,y)=\begin{cases}\frac{1}{\alpha}x^{\alpha}, & \text{if } y=0\\ \frac{1}{\alpha}x^{\alpha}+\delta\frac{1}{\alpha}y^{\alpha}-C, & \text{if } y>0\end{cases}$ where $\alpha<1$ and $C>0$	Constant	Constant if $y>0$	Yes, and is less pronounced at a higher stake
A simple version of Holden and Quiggin's (2017) zooming model: $U(x,y)=\frac{\left(x+b(x+y)^{\beta}\right)^{\alpha}}{\alpha}+\delta\frac{\left(y+b(x+y)^{\beta}\right)^{\alpha}}{\alpha}$ where $0<\beta<\alpha<1$ and $b>0$	Decreasing in the stake, and converging to $\frac{1}{1-\alpha}$	Decreasing from 1 to δ	No

3.4 Measuring Preferences and Testing Models: Experiment

3.4.1 Design

We implemented our measurement method in a real-stake experiment. Specifically, we measure subjects' preferences over intertemporal profiles consisting of two monetary rewards, one received in $t_1=1$ day, and the other received in $t_2=141$ days. The two dates are both in the future, and thus the present bias is ruled out.

For each subject, we elicit two indifference curves in the (x,y) space, one through the profile (€20, €20), and the other through (€80, €80). For each indifference curve, we use choice lists to elicit the indifference relation between the balanced profile and unbalanced profiles whose income ratio is among the following four values: 0, $\frac{1}{2}$, 2 or ∞. Hence, the following eight indifference relations are elicited from each subject:

$\left(0,y_{1,L}\right)\sim(20,20),\left(x_{2,L},2x_{2,L}\right)\sim(20,20),\left(2y_{3,L},y_{3,L}\right)\sim(20,20),\left(x_{4,L},0\right)\sim(20,20),$

$(0, y_{1,H}) \sim (80,80), (x_{2,H}, 2x_{2,H}) \sim (80,80), (2y_{3,H}, y_{3,H}) \sim (80,80), (x_{4,H}, 0) \sim (80,80)$. Figure 3.6 presents those indifference relations in the (x, y) space, as OA, OB, OC and OD for the lower stake, and O'A', O'B', O'C' and O'D' for the higher stake. Those indifference relations imply eight ARSs, denoted by $r_{j,k}$, where $j = 1,2,3,4, k = L, H$.

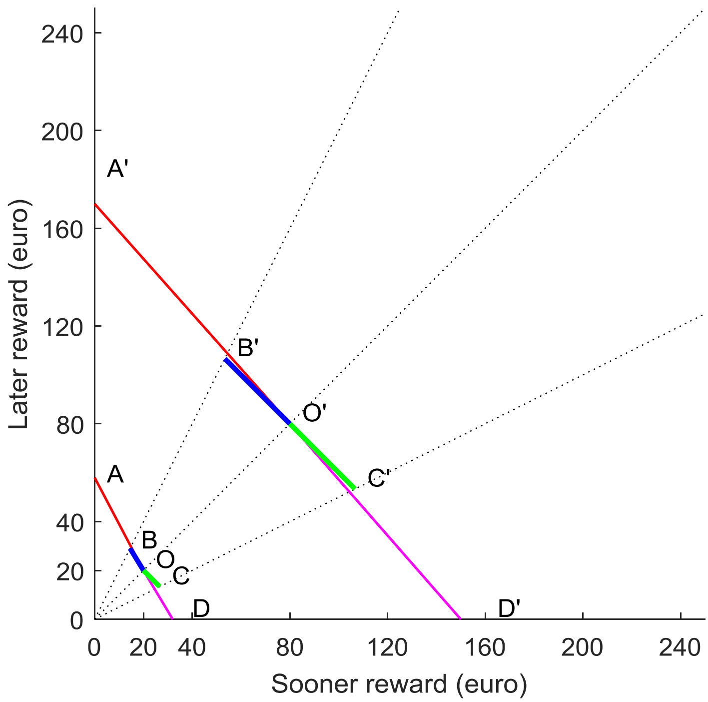


Figure 3.6: Testing Hypotheses by Comparing ARSs and Ratios of ARSs

The eight ARSs have variation in three aspects. First, half of the ARSs are for the lower stake (when $k = L$), and the other half for the higher stake (when $k = H$). Therefore, it is possible to identify the magnitude effect. Second, half of the ARSs are between a balanced profile and an unbalanced one in the interior (when $j = 2,3$), and the other half are between a balanced profile and a single dated reward (when $j = 1,4$). Hence, it is possible to test the all-sooner effect, and to test the overall convexity and the convexity in the interior, respectively. Third, half of the ARSs are in the future-leaning domain (when $j = 1{,}2$), and the other half in the present-leaning domain (when $j = 3{,}4$). Thus, the GDF and EIS in the interior can be measured.

Each choice list consists of 30 rows. Each of the four lists in the future-leaning domain (i.e. $j = 1{,}2$ and $k = L, H$) starts with a row implying an ARS a bit lower than 0.5 and ends with a row with a dominated option. The implied ARSs of the rows in between are increasing, and are the same in the four lists. In contrast, each of the four lists in the present-leaning domain (i.e. $j = 3{,}4$ and $k = L, H$) starts with a row implying an ARS a bit higher than 2 and ends with a row with a dominated option. The implied ARSs of the rows in between are decreasing, and are the same in the four lists. The lists in the future-leaning domain and those in the present-leaning domain are symmetric to the diagonal of the space. (See Appendix 3.C for all parameters used in the eight lists.) Figure 3.7 displays the 240 choice problems in the (x, y) space. Each choice problem is represented by a colored line segment, of which the two endpoints stand for the two options.

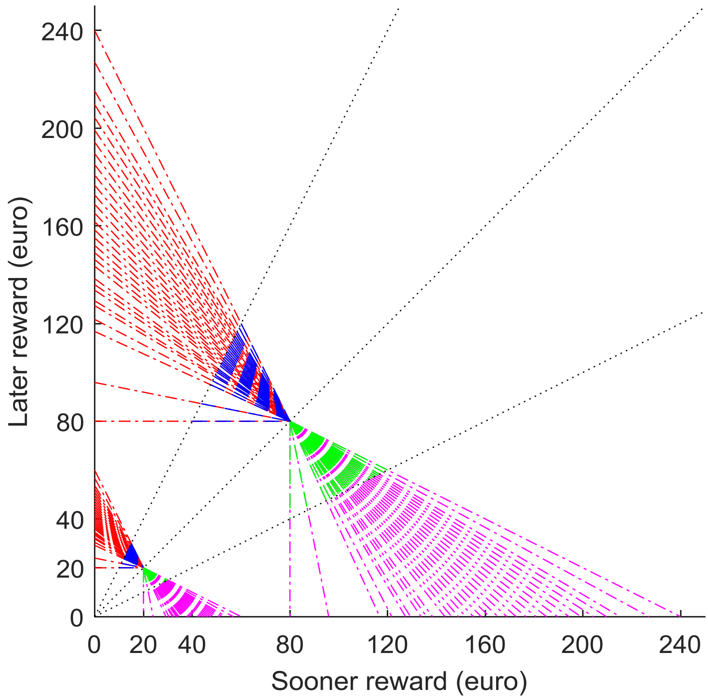


Figure 3.7: Binary Choice Problems Used in the Experiment

The comparability between the four lists within a domain (future-leaning or present-leaning) and the symmetry between the two domains make it possible to test the all-sooner effect and the magnitude effect simply by comparing the ARSs identified from the choice lists.

In principle, subjects need to make 30 choices in each list. However, we explicitly point out to subjects that in each list, one option is fixed and the other is getting worse as one moves down the list. Thereby, once a subject chooses the varying option in one row, we point out that she should also prefer the varying option in all preceding rows, and the program automatically selects the varying option in those rows. We do the same thing for choosing the fixed option. Figure 3.2 displays what happens if a subject chooses an option in a row. By imposing such a mechanism, we rule out the possibility of multiple switching.

The order of the eight lists are randomized in two aspects: from the sooner boundary to the later boundary or vice versa, and from the lower stake to the higher stake or vice versa. Subjects can switch between lists at any moment, regardless of whether the current list is finished. Choices are automatically stored when one switches to another list. Therefore, subjects can easily make comparisons across lists if they like. Choices can be submitted only if all the 240 decisions in the eight lists are completed.

The experimental payments have two components. First, all subjects receive a €3 *participation fee* on each of the two dates, t_1 and t_2. Second, each subject has a 10 percent chance to receive *earnings depending on choices*.

The earnings depending on choices are determined by a random incentive scheme with both between-subject and within-subject randomization. Prior to the decision-making stage, each subject is randomly given a lottery number, ranging from 0 to 9. At the end of each session, one of the subjects throws a ten-sided die in front of all subjects in that session. Subjects who have a lottery number identical to the die roll receive the earnings depending on choices. For each of them, one choice is randomly selected to be realized, and the amounts of money in that choice are added to the participation fees on the two dates, respectively. The between-subject randomization makes it possible to use high-stake rewards, and hence the magnitude effect can be investigated. The within-subject randomization makes all choices incentivized and it helps avoid wealth effect and hedging.

For a real-stake experiment on intertemporal choices, it is important to equalize the transaction costs as well as the credibility of payments across periods. For this purpose, we provide an equal amount of participation fees on both the payment dates, and the participation fees and the earnings depending on choices are both delivered by bank transfer. Thus, the payment tools are the same for the two payments, and subjects need to receive money twice regardless of their choices. A survey in Sun and Potters (2016) shows that bank transfer is as good as cash in terms of liquidity to Dutch university students. Therefore, subjects have reasons to believe that they will receive the payments on time and then the money can be used immediately. Moreover, we provide each subject a payment reminder card, which records the amounts and the dates of the payments, as well as the contact information of the experimenter. This should further increase the credibility of the payments.

The experiment was conducted at the CentERlab, Tilburg University in September of 2017.[26] 114 students of the university participated in one of the eight sessions. One session took an hour on average. At the start of each session, the experimenter read the instructions aloud in front of all subjects. Then subjects made choices in a zTree program (Fischbacher, 2007). 12 subjects got the earnings depending on decisions, which averaged €82.46. The overall average earning was €14.68.

3.4.2 *Analyses*

We address our three questions based on the eight ARSs, $r_{j,k}, j = 1,2,3,4, k = L, H$.

The first question is whether the EIS is larger for higher stakes. Formally, the EIS is larger for the higher stake than for the lower stake if

$$\frac{r_{2,L}}{r_{3,L}} < \frac{r_{2,H}}{r_{3,H}}. \tag{3.10}$$

In Figure 3.6, $-\frac{1}{r_{2,L}}$ and $-\frac{1}{r_{3,L}}$ correspond to the slope of OB and that of OC, respectively, and $-\frac{1}{r_{2,H}}$ and $-\frac{1}{r_{3,H}}$ correspond to the slope of O'B' and that of O'C', respectively. Then the EIS is larger for the higher stake if the indifference curve between B' and C' is less convex than that between B and C.

The second question is whether the GDF is larger for a higher stake. Formally, the GDF is larger for the higher stake than for the lower stake if

[26] All payment dates are workdays, which guarantees the punctual arrival of payments via bank transfer.

(3.11) $$\sqrt{r_{2,L} r_{3,L}} < \sqrt{r_{2,H} r_{3,H}}.$$

Here, $\sqrt{r_{2,L} r_{3,L}}$ is a local estimate of the GDF for the lower stake. In Figure 3.6, the GDF is larger for the higher stake than for the lower stake if the curve B'C' is on average flatter than the curve BC.

The third question concerns the existence of the all-sooner effect. The intuition is that people are less patient when they are choosing between a pure sooner reward and a mixed profile than between two mixed profiles. Formally, the all-sooner effect exists when the stake is low if

(3.12) $$r_{4,L} < r_{3,L}.$$

In Figure 3.6, $-\frac{1}{r_{4,L}}$ and $-\frac{1}{r_{3,L}}$ correspond to the slope of OD and that of OC, respectively. If OD is steeper than OC, the all-sooner effect exists when the stake is low. The same test applies to the indifference curve with the higher stake. If O'D' is steeper than O'C', the all-sooner effect exists when the stake is high.

3.5 Results

Each of the 114 subjects made 240 choices in eight choice lists. Two subjects never switched in any list, and their choices are not consistent across lists. They apparently did not understand the task, and hence we remove their choices from the sample. Five other subjects chose a dominated option at least once. We also rule out their choices.

Among the remaining 107 subjects, 24 of them (22 percent) always behave as if they maximize the total amount without discounting.[27] Those subjects are likely to be people who fully integrate experimental rewards with their lifelong

[27] Since the minimum positive annual rate of return in our design is 11.6 percent, which is much higher than the market interest rate, a subject who discounts money exponentially with the market interest rate will also make choices as if they are maximizing the total amount. One concern about the measurement method is that the specific design might have induced some subjects to use the heuristic of choosing the option with the largest total amount. Unfortunately, we cannot distinguish people who are induced to be total-amount-maximizers from those who are truly (approximately) total-amount-maximizers. But given the fraction being 22%, which is much smaller than 38% in Andreoni and Sprenger (2012a), we expect the problem to be less severe than the CTB method (Andreoni and Sprenger, 2012a).

wealth. Since those people present no magnitude effect nor all-sooner effect, we focus on the remaining 83 subjects throughout this section. All tests we performed are unaffected by the exclusion of total-amount maximizers.

Given the often-observed heterogeneity in preferences, we do not "force" all subjects to have the same preference and estimate the "representative" preference. Instead, we perform within-subject comparisons (e.g. between the slope of OC and that of OD) so that the three questions are addressed in a way which allows subjects to have different preferences (e.g. OC and OD can be different across subjects).[28]

Throughout we treat censored observations as extreme values (i.e. zero or infinity ARS). Since only a small fraction of choices are censored (5 percent) and the signed rank test only relies on the rank, we expect censoring has limited effect on our results.

3.5.1 Comparability with Previous Studies

Before we address our main questions, we first check the validity of our dataset by examining if several stylized facts can be observed in our dataset.

One pattern established in the literature is that people are indifferent between a smaller pure sooner reward and a larger pure later reward (see Frederick et al. 2002 for a review). In other words, the monetary discount factor is less than one.[29] Formally,

$$\hat{\delta}_k^m \equiv \left(\frac{x_{4,k}}{y_{1,k}}\right)^{\frac{365}{140}} < 1, \qquad k = L, H. \tag{3.13}$$

[28] We perform sign tests and Wilcoxon signed rank tests. The sign test has the advantage of allowing heterogeneity in the size of the effect being tested. The Wilcoxon signed rank test has an advantage in power. The results of the two tests are always consistent in our study, suggesting that our findings are robust.

[29] In a statistical test, noise is taken into account. Thereby, strictly speaking, we are testing whether the monetary discount factor is more likely to be less than one than to be greater than one. Similar arguments apply to all other tests in this paper.

We thus test it in our dataset Table 3.2 shows the monetary discount factors for the lower stake and for the higher stake, respectively. The medians are less than one, confirming that subjects exhibit impatience in choices between single dated rewards.

Table 3.2: Monetary Discount Factors between Single Dated Rewards

	Monetary discount factor for the lower stake: $\hat{\delta}_L^m = \left(\frac{x_{4,L}}{y_{1,L}}\right)^{\frac{365}{140}}$	Monetary discount factor for the higher stake: $\hat{\delta}_H^m = \left(\frac{x_{4,H}}{y_{1,H}}\right)^{\frac{365}{140}}$
Median	0.896	0.947
$\#(\hat{\delta}_k^m < 1)$	65	55
$\#(\hat{\delta}_k^m = 1)$	15	25
$\#(\hat{\delta}_k^m > 1)$	3	3
Sign test:	p = 0.000***	p = 0.000***
Wilcoxon signed rank test:	p = 0.000***	p = 0.000***

Notes: Two-sided sign tests and Wilcoxon signed rank tests on the logarithm of the monetary discount factors for the lower stake and for the higher stake, respectively. The null hypothesis is that the monetary discount factor is equal to one. Censored observations are treated as extreme values (zero or infinity ARS). ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

The second stylized fact is the magnitude effect over single dated rewards: people have a larger monetary discount factor for higher stakes. Formally, it means

$$\hat{\delta}_L^m < \hat{\delta}_H^m. \tag{3.14}$$

We also test it with our data, and the results are presented in Table 3.3. The median ratio of the monetary discount factors for the higher stake to that for the lower stake is greater than one. Thus, the magnitude effect on the monetary discount rate is significant, which is consistent with previous evidence.

Third, in their parametric estimation with a classic discounting model, Andreoni and Sprenger (2012a) obtained a discount factor less than one. Since the discount factor in a classic discounting model corresponds to the GDF in a general case, we examine if the GDF in our dataset is on average less than one. Table 3.4 displays the estimates of the GDFs at the two magnitudes. The medians at both magnitudes are less than one, which is consistent with the previous finding.

Table 3.3: Magnitude Effect on the Monetary Discount Factor

	Ratio of the monetary discount factor between the two magnitudes: $\frac{\hat{\delta}_H^m}{\hat{\delta}_L^m}$
Median	1.057
$\#\left(\frac{\hat{\delta}_H^m}{\hat{\delta}_L^m} > 1\right)$	45
$\#\left(\frac{\hat{\delta}_H^m}{\hat{\delta}_L^m} = 1\right)$	14
$\#\left(\frac{\hat{\delta}_H^m}{\hat{\delta}_L^m} < 1\right)$	21
Sign test:	p = 0.004***
Wilcoxon signed rank test:	p = 0.000***

Notes: Two-sided sign test and Wilcoxon signed rank test on the logarithm of the ratio of the monetary discount factor for the higher stake to that for the lower stake. The null hypothesis is that the monetary discount factors are equal for the two stakes. Censored observations are treated as extreme values (zero or infinity ARS), and indeterminate forms are dropped. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

Table 3.4: Generalized Discount Factors

	Generalized discount factor for the lower stake: $\hat{\delta}_L = \left(\sqrt{\hat{r}_{2,L}\hat{r}_{3,L}}\right)^{\frac{365}{140}}$	Generalized discount factor for the higher stake: $\hat{\delta}_H = \left(\sqrt{\hat{r}_{2,H}\hat{r}_{3,H}}\right)^{\frac{365}{140}}$
Median	0.896	0.947
$\#(\hat{\delta}_k < 1)$	48	51
$\#(\hat{\delta}_k = 1)$	29	23
$\#(\hat{\delta}_k > 1)$	6	9
Sign test:	p = 0.000***	p = 0.000***
Wilcoxon signed rank test:	p = 0.000***	p = 0.000***

Notes: Two-sided sign tests and Wilcoxon signed rank tests on the logarithm of the generalized discount factors for the lower stake and for the higher stake, respectively. The GDFs are estimated as the geometric mean of the ARSs in the interior. The null hypothesis is that the GDF is equal to one. Censored observations are treated as extreme values (zero or infinity ARS). ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

Fourth, experiments using intertemporal profiles found that the interior part of an indifference curve is convex (Cheung, 2015a) or the entire indifference curve is on average convex (Andreoni and Sprenger, 2012a; Sun and Potters, 2016). We thereby test if it is also the case in our dataset. Table 3.5 shows the tests on convexity. The second and third columns are about the convexity of the lower-stake indifference curve. The second column presents the ratio between the two ARSs in the interior, which is a measure of convexity in the interior. Given that the ratio is on average greater than one, the interior part of the lower-stake indifference curve is convex. The third column presents the ratio of the two ARSs between the balanced profile and the two boundaries, which is a measure of average convexity for the entire indifference curve. Since the ratio is greater than one, the lower-stake indifference curve is on average convex. The fourth and fifth columns show the higher-stake case. The indifference curve in the interior is still convex, but no evidence shows that the entire curve on average is convex. The fact that the indifference curve with the higher stake is on average linear is consistent with the results from the parametric estimation by Sun and Potters (2016).

Table 3.5: Convexity of Indifference Curves in the Interior and Overall

Ratios of two ARSs as measures of convexity	For the lower stake		For the higher stake	
	Interior: $\frac{\hat{r}_{3,L}}{\hat{r}_{2,L}}$	Overall: $\frac{\hat{r}_{4,L}}{\hat{r}_{1,L}}$	Interior: $\frac{\hat{r}_{3,H}}{\hat{r}_{2,H}}$	Overall: $\frac{\hat{r}_{4,H}}{\hat{r}_{1,H}}$
Median	1.043	1.043	1.000	1.000
$\#\left(\frac{\hat{r}_{j_2,k}}{\hat{r}_{j_1,k}} > 1\right)$	48	42	40	32
$\#\left(\frac{\hat{r}_{j_2,k}}{\hat{r}_{j_1,k}} = 1\right)$	21	24	23	27
$\#\left(\frac{\hat{r}_{j_2,k}}{\hat{r}_{j_1,k}} < 1\right)$	14	17	20	24
Sign test:	0.000***	0.002***	0.013**	0.350
Wilcoxon signed rank test:	0.000***	0.008***	0.034**	0.534

Notes: Ratios of two ARSs in the interior as a measure of convexity in the interior, and ratios of two ARSs between a balanced profile and the two boundaries as a measure of average convexity of the entire indifference curve. The indifference curve is convex if the ratio is greater than one. Two-sided sign tests and Wilcoxon signed rank tests on the logarithm of the measures. The null hypothesis is that the measurement of convexity is equal to one, meaning that the indifference curve is linear. Censored observations are treated as extreme values (zero or infinity ARS). ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

Given that all the results replicate the stylized facts found in the literature, we have reason to believe that our design did not bias preferences in a systematic way.

3.5.2 *Main Results*

In this subsection, we address the three questions regarding the two channels of the magnitude effect and the all-sooner effect.

First, we test if the EIS is larger for higher stakes. The EIS is measured by the ratio of the two ARSs in the interior. The larger the measure, the more convex the indifference curve is, and the smaller the EIS is. Figure 3.8 shows the distributions of the measures of convexity for the lower stake and for the higher stake, respectively. The measure is smaller for the higher stake, suggesting that the EIS is larger for the higher stake. We then perform formal tests on the measures for the two different stakes. Table 3.6 presents the results of the tests. The ratio between the measures of convexity is less than one, implying that the indifference curve is less convex for the higher stake, or equivalently, the EIS is larger for the higher stake.

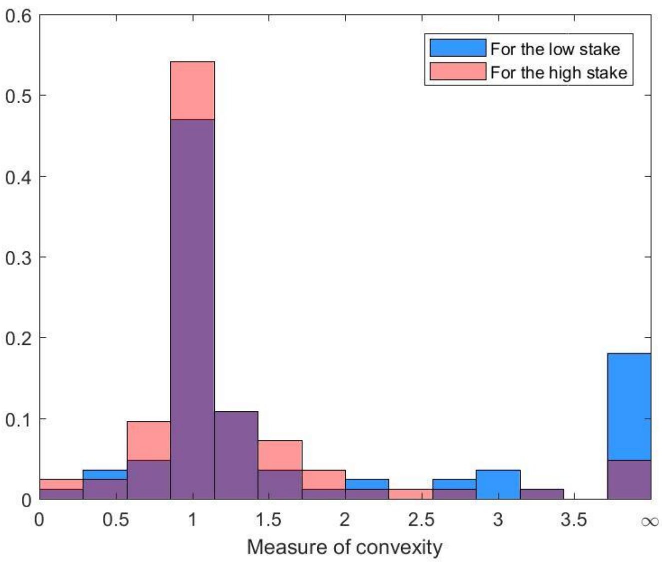


Figure 3.8: Distributions of the Measures of Convexity for the Lower Stake and for the Higher Stake

Table 3.6: Magnitude Effect on the Elasticities of Intertemporal Substitution

Ratio of the measures of convexity between the two stakes	$\frac{\hat{r}_{3,H}}{\hat{r}_{2,H}}\frac{\hat{r}_{2,L}}{\hat{r}_{3,L}}$
Median	0.959
$\#\left(\frac{\hat{r}_{3,H}}{\hat{r}_{2,H}}\frac{\hat{r}_{2,L}}{\hat{r}_{3,L}} < 1\right)$	43
$\#\left(\frac{\hat{r}_{3,H}}{\hat{r}_{2,H}}\frac{\hat{r}_{2,L}}{\hat{r}_{3,L}} = 1\right)$	17
$\#\left(\frac{\hat{r}_{3,H}}{\hat{r}_{2,H}}\frac{\hat{r}_{2,L}}{\hat{r}_{3,L}} > 1\right)$	23
Sign test:	0.019**
Wilcoxon signed rank test:	0.010**

Notes: Two-sided sign test and Wilcoxon signed rank test on the logarithm of the ratio of the measure of convexity for the higher stake to that for the lower stake. The EIS is larger for a higher stake if the ratio is less than one. The null hypothesis is that the ratio is equal to one, meaning that the indifference curve for the lower stake and that for the higher stake are equally convex. Censored observations are treated as extreme values (zero or infinity ARS). ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

Second, we test if the GDF is also larger for higher stakes. The GDF is measured by the geometric average of the two ARSs in the interior. Figure 3.9 shows the distributions of the GDFs for the lower stake and for the higher stake, respectively. The GDF seems to be larger for the higher stake, but it is not clear if it is significant. Table 3.7 presents the result of the test. The ratio of the GDF for the higher stake to that at the lower stake is not significantly greater than one.[30] Thereby, we find no evidence that the GDF is larger for higher stakes.

[30] We perform a power analysis. With the alternative hypothesis that the proportion is 0.3, the rejection rate is 92%. Thus, our test is not underpowered.

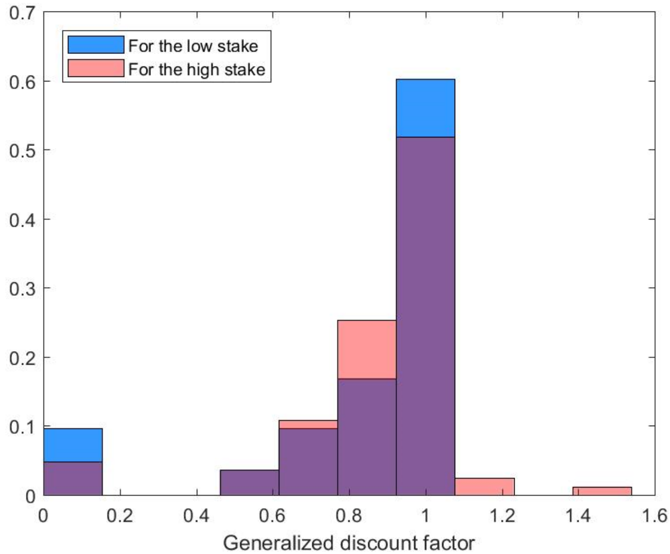


Figure 3.9: Distributions of the Generalized Discount Factors for the Lower Stake and for the Higher Stake

Table 3.7: Magnitude Effect on the Generalized Discount Factor

Ratio of the GDF between the two stakes:	$\frac{\hat{\delta}_H}{\hat{\delta}_L}$
Median	1.0000
$\#\left(\frac{\hat{\delta}_H}{\hat{\delta}_L} > 1\right)$	36
$\#\left(\frac{\hat{\delta}_H}{\hat{\delta}_L} = 1\right)$	22
$\#\left(\frac{\hat{\delta}_H}{\hat{\delta}_L} < 1\right)$	24
Sign test:	p = 0.155
Wilcoxon signed rank test:	p = 0.134

Notes: Two-sided sign test and Wilcoxon signed rank test on the logarithm of the ratio of the GDF for the higher stake to that for the lower stake. The null hypothesis is that the ratio is equal to one. Censored observations are treated as extreme values (zero or infinity ARS), and indeterminate forms are dropped. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

From Figure 3.9, we also notice that there are some subjects who have an ARS of zero for the lower stake but have an ARS around one for the higher stake. In other words, they are extremely impatient when the stake is lower, but patient when the stake is higher. This is consistent with the magnitude-dependent discounting model by Noor (2011). The existence of extreme subjects suggests that there is heterogeneity in how GDF changes with the stake.

The results on the two channels of the magnitude effect imply that the magnitude of outcomes affects choices mainly through the perceived fungibility of rewards across periods. In other words, when the magnitude is larger, people care more about the discounted sum of amounts and less about smoothing. On the other hand, there is no evidence showing that the general patience is higher when the magnitude is larger, though heterogeneity seems to exist.

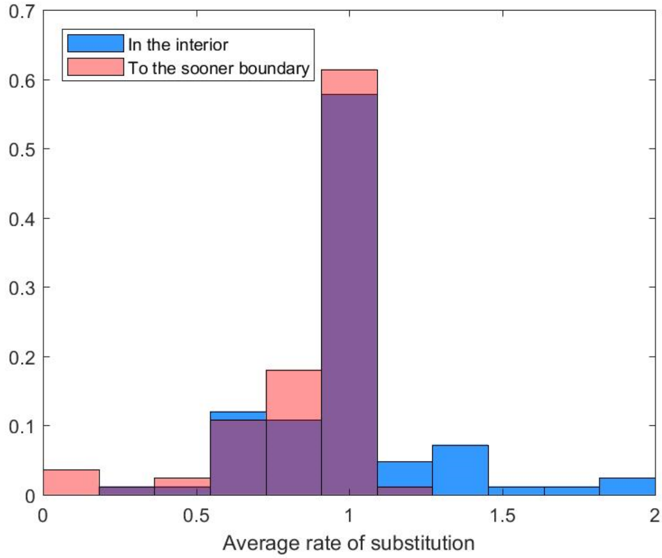
Our results about the magnitude effect have implications for estimating preferences. If researchers estimate preferences with low-stake choices, omitting the magnitude effect leads to an overestimation of the utility curvature.

Third, we look at the all-sooner effect. The all-sooner effect exists if people are less patient when choosing between a pure sooner reward and a mixed profile than when choosing between two mixed profiles, i.e. if the ARS to the sooner boundary is less than that in the interior. Figure 3.10(a) shows the distributions of the two ARSs for the lower stake. The ARS to the sooner boundary is on average less than that in the interior, suggesting the existence of the all-sooner effect. Figure 3.10(b) shows the distributions of the two ARSs for the higher stake. The two ARSs are about the same. We then perform tests on the ratios of the two ARSs to check if they are different from one. Table 3.8 shows the results of the tests. For the lower stake, most subjects have a smaller ARS to the sooner boundary than in the interior, which implies the existence of the all-sooner effect. For the higher stake, the difference between the two ARSs are smaller. This is consistent with the fixed-cost-of-waiting model by Benhabib et al. (2010) because a fixed cost of waiting implies that the extra disutility of a mixed profile than a pure sooner reward is less pronounced for higher stakes, and hence people exhibit less extra impatience when a pure sooner reward is available.

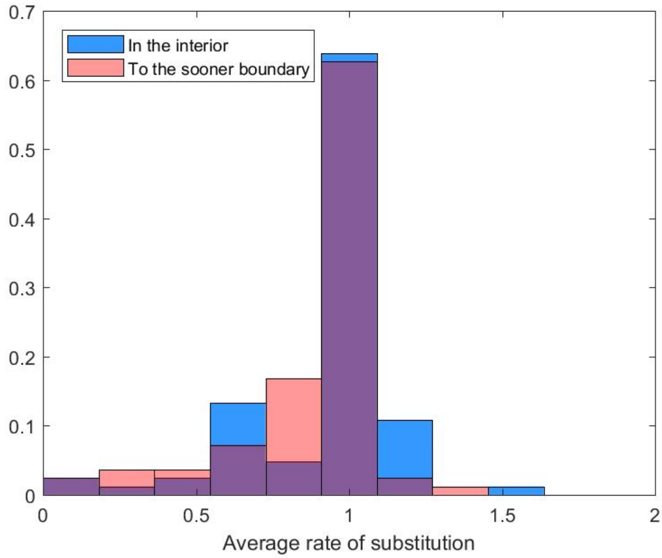
The existence of the all-sooner effect also has implications for both theoretical and empirical studies. First, it implies that impatience concerns more than discounting. Most of the impatience observed in choices between single dated rewards is probably caused both by a proportional cost of delaying rewards (discounting) and by a disproportional cost (the all-sooner effect). The two channels have different effects on choices between general intertemporal profiles. Second, the all-sooner effect implies a violation of convexity near the sooner boundary. The intuition is that people have a preference for pure sooner rewards, which dominates the preference for smoothing when the degree of smoothing is low anyway (since almost all rewards are on the sooner date). Third, if researchers estimate preferences with some single dated rewards, omitting the all-sooner effect leads to an overestimation of the discount rate, because part of the impatience is actually due to the all-sooner effect. In the meantime, it leads to an underestimation of the utility curvature because the average curvature is lowered by the non-convexity resulted from the all-sooner effect.

Our evidence of the all-sooner effect combined with our evidence of convexity shows that existing evidence of convexity in the interior and overall is not conflicting with the all-sooner effect; the all-sooner effect only leads to a violation of convexity near the sooner boundary, while in the interior and overall indifference curves are still convex in our dataset.

The result that we do not find a magnitude effect on the generalized discount factor does not contradict the previous evidence of the magnitude effect on the discount rate in Sun and Potters (2016). In Sun and Potters (2016), the discount rate is measured parametrically without considering the all-sooner effect, and hence it could be a mixture of the real discount rate and the all-sooner effect. It is possible that the discount rate does not change with stakes while the all-sooner effect is less salient when outcomes are larger. Our result of the all-sooner effect is consistent with this story.



(a) For the lower stake



(b) For the higher stake

Figure 3.10: Distributions of the Average Rates of Substitution in the Interior and to the Sooner Boundary

Table 3.8: The All-Sooner Effect

Ratios of the ARS to the sooner boundary to that in the interior as measures of the all-sooner effect	For the lower stake: $\frac{\hat{r}_{4,L}}{\hat{r}_{3,L}}$	For the higher stake: $\frac{\hat{r}_{4,H}}{\hat{r}_{3,H}}$
Median	0.959	1.000
$\#\left(\frac{\hat{r}_{4,k}}{\hat{r}_{3,k}} < 1\right)$	43	35
$\#\left(\frac{\hat{r}_{4,k}}{\hat{r}_{3,k}} = 1\right)$	25	27
$\#\left(\frac{\hat{r}_{4,k}}{\hat{r}_{3,k}} > 1\right)$	15	20
Sign test:	0.000***	0.058*
Wilcoxon signed rank test:	0.000***	0.039**

Notes: Two-sided sign tests and Wilcoxon signed rank tests on the logarithm of the ratio of the ARSs in the interior to that to the sooner boundary. The null hypothesis is that the ratio is equal to one. The all-sooner effect exists if the ratio is less than one. Censored observations are treated as extreme values (zero or infinity ARS), and indeterminate forms are dropped. ***, ** and * indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

We compare our results with the predictions of existing models considered in Table 3.1. First, given that we find strong evidence of the all-sooner effect, the fixed-cost-of-waiting model has an advantage in predicting such a behavior. If people behave as if they incur a fixed cost whenever they forgo a pure sooner reward, they appear less patient when a pure sooner reward is available. Second, only one model predicts an increase in the EIS with the stake: the mental accounting version of Fudenberg and Levine's (2006) bank-nightclub model. The intuition is that people perceive small rewards as extra consumption in the short-run, which is less fungible, while they perceive large rewards primarily as contributions to life-long savings, which is highly fungible. Third, since we do not find evidence that the GDF is larger for higher stakes, most models considered are consistent with our results in this term. As mentioned above, for subjects who are extremely impatient for the lower stake but patient for the higher stake, the magnitude-dependent discounting model by Noor (2011) can give an explanation. Nevertheless, none of the models under consideration gives a uniform account for the empirical patterns in all of the three aspects.

3.5.3 A Simple and Flexible Model that Captures All Findings

We propose a simple model which can accommodate all the characteristics found in our results. The model is

$$U(x,y)=\begin{cases}(1-\mu)x+\mu\dfrac{(x+\omega)^{\alpha}}{\alpha}+\delta_0\mu\dfrac{\omega^{\alpha}}{\alpha}, & \text{if } y=0\\(1-\mu)x+\mu\dfrac{(x+\omega)^{\alpha}}{\alpha}+\delta_{\infty}(1-\mu)y+\delta_0\mu\dfrac{(y+\omega)^{\alpha}}{\alpha}-C, & \text{if } y>0\end{cases}\tag{3.15}$$

where α is the utility curvature parameter satisfying $\alpha<1$, δ_0 and δ_∞ are two discount factors satisfying $0\le\delta_0\le\delta_\infty\le 1$, C is a fixed cost of waiting satisfying $C>0$, and μ is a weight satisfying $\mu\in[0,1]$.

The model has an intuitive interpretation: the utility function in each period consists of a linear part and a concave part, representing the goals of life-long wealth maximization and short-run utility maximization, respectively. μ is the weight of the myopic motivation. As the stake increases, the concave part becomes relatively less important, and hence the sooner reward and the later reward become more fungible (i.e. the EIS increases). On the other hand, the short-run part is discounted more than the long-run part ($\delta_0\le\delta_\infty$), so that when the stake increases, the effective weight of the later reward becomes higher (i.e. the GDF increases).

The model represents a preference with the following characteristics: i) it has convex indifference curves in the interior; ii) the all-sooner effect exists, and its effect on the ARS is decreasing; iii) the EIS first decreases and then increases to ∞ as the stake increases from 0 to ∞; iv) the GDF changes monotonically from a value between δ_0 and δ_∞ to δ_∞ as the stake increases from 0 to ∞. (See Appendix 3.B for a proof.)

The model is flexible to accommodate different types of decision makers. By assuming $C = 0$, the all-sooner effect disappears. By further assuming $\mu = 0$, it turns into a discounting model with linear utility functions, and hence can represent the preference of total-amount maximizers. With $\delta_0 = \delta_\infty$, the model represents the preference of people whose GDF is constant. With $C = 0$, $\mu = 1$ and $\omega = 0$, the model represents the preference of people whose EIS is constant.

Lastly, the model is tractable, with simple functional forms of utilities and marginal utilities, and hence can be easily estimated in empirical studies.

Given the consistency with empirical findings, the flexibility and the tractability, the model can be used to estimate preferences and predict choices in future empirical studies, especially in those involve both single dated rewards and mixed profiles and have variation in stakes.

3.6 Conclusion and Discussions

We develop a simple method for measuring intertemporal preferences. The method is parameter-free, incentive-compatible and independent of time horizon effects. It requires weak assumptions on the preferences to be measured, and hence it can be used to study a variety of models.

We apply the method to measure preferences over two-outcome intertemporal profiles. The measurement is used for testing a few existing models of intertemporal choices in a parameter-free way. Those models vary in their predictions on how choices between intertemporal profiles change with the stake (the magnitude effect) and whether people have a preference for pure sooner rewards over mixed profiles (the all-sooner effect). We find three predominant characteristics. First, the elasticity of intertemporal substitution is larger for higher stakes, meaning that people feel monetary rewards are more fungible across time when the magnitude of the rewards is larger. Second, we do not find evidence that the generalized discount factor is larger for higher stakes, suggesting that on average discounting is not affected by the magnitude of rewards. Third, we find evidence of the all-sooner effect, meaning that people have a preference for pure sooner rewards.

Our results provide implications for the practice of preference elicitation in scientific studies and prescriptive consultation. Our finding about the magnitude effect suggests that estimating preferences with low-stake tasks leads to an overestimation of the utility curvature if the magnitude effect is omitted. Our finding of the all-sooner effect suggests that measuring preferences using choices between single dated rewards without considering the all-sooner effect leads to an overestimation of the discount rate and an underestimation of the utility curvature. The simple model we provide can be used to correct for such biases in empirical work where both single dated rewards and intertemporal profiles are involved, or choices of various magnitudes are included.

Although we fix the time points of the two outcomes and focus on how choices are affected by changes in outcomes in this study, our measurement method can be used when time points are varied. By measuring discount factors for different front-end delays and delays, it is possible to measure the entire discount function. Meanwhile, it is easy to combine our empirical findings with the findings about the time horizon effects. For instance, in our simple model, δ_0 and δ_∞ can be functions decreasing in τ if the discount function is only sensitive in the delay (Read, 2001; Dohmen et al., 2017). If the front-end delay also matters, δ_0 and δ_∞ should also be applied to the first-period utility function, and all discount functions can be decreasing in the time distance from now.

In this study we only measure preferences over two-outcome profiles, but our framework can be extended to profiles with any number of periods. In that case, the EIS and the GDF can be defined for any two consecutive periods. Are all the EISs increasing in the stake? Does the all-sooner effect occur whenever the choice is between a single dated reward at t_i and a profile with rewards no earlier than t_i? Those are interesting questions that can be addressed in further studies.

A last observation concerns the non-monotonicity caused by the assumption of a fixed cost of waiting. The preference represented by the utility function is discontinuous and non-monotonic between the sooner boundary and the rest of the space. As a result, a pure sooner reward might be preferred to a mixed profile which dominates the former one. For instance, suppose $U(30,0) = U(34,2)$, which means the pure sooner reward $(30,0)$ is as good as the profile $(34,2)$. But this implies that $(31,0)$ is preferred to $(34,2)$ even though $(31,0)$ is dominated by $(34,2)$ in quantity. By a thought experiment, it is hard to believe that subjects will ever choose the dominated option. This problem matters when choices with dominated options are involved. One way to solve this problem is to assume the preference to be "$(x_1, y_1) \succcurlyeq (x_2, y_2)$ if and only if, either $x_1 \geq x_2$ and $y_1 \geq y_2$, or $(x_1 - x_2)(y_1 - y_2) < 0$ and $U(x_1, y_1) \geq U(x_2, y_2)$". In other words, a choice is made first by checking if one option dominates the other. Only if no domination exists, is the choice determined by the utility function. This decision-making rule, however, implies cyclicity. For instance, if $U(31,0) > U(30,0) = U(20,20) = U(34,2) > U(33,1)$, it holds that $(31,0) \succ (20,20)$, $(20,20) \succ (33,1)$ and $(33,1) \succ (31,0)$. Given that cyclicity is observed in intertemporal choices (consider the subadditivity in discounting. See Read, 2001), the intransitive model should not be rejected without empirical investigation, but ought to be tested in a future study.

Appendix to Chapter 3

3.A Proofs of the Predictions of Various Existing Models

For a general intertemporal utility function, $U(x, y)$, the EIS at (x, y) is

$$\varepsilon(x, y) \equiv \frac{1}{\frac{\partial}{\partial Q} \ln \rho\big(x(Q, u), y(Q, u)\big)},$$

where $Q \equiv \ln\left(\frac{x}{y}\right)$ and $u = U(x, y)$.

If the intertemporal utility function is additively separable, it can be written as

$$U(x, y) = u_1(x) + u_2(y),$$

where $u_1(\cdot)$ and $u_2(\cdot)$ are two period utility functions. The GDF in this case is

$$\delta(y) = \frac{u_2'(y)}{u_1'(y)}.$$

The EIS is

$$\varepsilon(x, y) = -\frac{u_1'(x)u_2'(y)[u_1'(x)x + u_2'(y)y]}{\left[u_2''(y)\big(u_1'(x)\big)^2 + u_1''(x)\big(u_2'(y)\big)^2\right]xy}$$

$$= \left[-\frac{u_2''(y)y}{u_2'(y)} \cdot \frac{u_1'(x)x}{u_1'(x)x + u_2'(y)y} - \frac{u_1''(x)x}{u_1'(x)} \cdot \frac{u_2'(y)y}{u_1'(x)x + u_2'(y)y}\right]^{-1}.$$

If further, the preference is represented by a classic discounting model, the intertemporal utility function can be written as

$$U(x, y) = u(x) + \delta u(y).$$

Then the GDF is

$$\delta(y) = \delta,$$

and the EIS is

$$\varepsilon(x, y) = \left[-\frac{u''(y)y}{u'(y)} \frac{u'(x)x}{u'(x)x + \delta u'(y)y} - \frac{u''(x)x}{u'(x)} \frac{\delta u'(y)y}{u'(x)x + \delta u'(y)y}\right]^{-1}.$$

3.A.1 Power Utility Function and Magnitude-Independent Discount Factor

The preference is represented by the intertemporal utility function

$$U(x,y)=\frac{1}{\alpha}x^{\alpha}+\delta\frac{1}{\alpha}y^{\alpha}$$

where $\alpha\leq 1$. The GDF is

$$\delta(y)=\delta,$$

and the EIS is

$$\varepsilon(x,y)=\frac{1}{1-\alpha}.$$

Hence, neither the GDF nor the EIS changes with the stake.

3.A.2 Power Utility Function with Positive Background Consumption and Magnitude-Independent Discount Factor

The preference is represented by the intertemporal utility function

$$U(x,y)=\frac{1}{\alpha}(x+\omega)^{\alpha}+\delta\frac{1}{\alpha}(y+\omega)^{\alpha}$$

where $\alpha\leq 1$ and $\omega>0$. The GDF is still

$$\delta(y)=\delta,$$

and the EIS is

$$\varepsilon(x,y)=\frac{1}{1-\alpha}\left[\frac{y}{y+\omega}\frac{(x+\omega)^{\alpha-1}x}{(x+\omega)^{\alpha-1}x+\delta(y+\omega)^{\alpha-1}y}\right.$$
$$\left.+\frac{x}{x+\omega}\frac{\delta(y+\omega)^{\alpha-1}y}{(x+\omega)^{\alpha-1}x+\delta(y+\omega)^{\alpha-1}y}\right]^{-1}.$$

Since $\frac{y}{y+\omega}$ and $\frac{x}{x+\omega}$ are both decreasing in the stake, and

$$\frac{(x+\omega)^{\alpha-1}x}{(x+\omega)^{\alpha-1}x+\delta(y+\omega)^{\alpha-1}y}+\frac{\delta(y+\omega)^{\alpha-1}y}{(x+\omega)^{\alpha-1}x+\delta(y+\omega)^{\alpha-1}y}=1,$$

the EIS is decreasing in the stake. For any $\frac{x}{y}=\lambda>0$,

$$\lim_{y\to 0} \varepsilon(x,y) = +\infty$$

and

$$\lim_{y\to +\infty} \varepsilon(x,y) = \frac{1}{1-\alpha},$$

the EIS decreases from $+\infty$ to $\frac{1}{1-\alpha}$ as the stake increases.

3.A.3 Exponential Utility Function and Magnitude-Independent Discount Factor

The preference is represented by the intertemporal utility function

$$U(x,y) = [1-\exp(-x)] + \delta[1-\exp(-y)].$$

The GDF is again

$$\delta(y) = \delta,$$

and the EIS is

$$\varepsilon(x,y) = \left[y\frac{x\exp(-x)}{x\exp(-x)+\delta y\exp(-y)} + x\frac{\delta y\exp(-y)}{x\exp(-x)+\delta y\exp(-y)} \right]^{-1}.$$

Since

$$\frac{x\exp(-x)}{x\exp(-x)+\delta y\exp(-y)} + \frac{\delta y\exp(-y)}{x\exp(-x)+\delta y\exp(-y)} = 1,$$

the EIS is decreasing in the stake. For any $\frac{x}{y} = \lambda > 0$,

$$\lim_{y\to 0} \varepsilon(x,y) = +\infty$$

and

$$\lim_{y\to +\infty} \varepsilon(x,y) = 0,$$

the EIS decreases from $+\infty$ to 0 as the stake increases.

3.A.4 Loewenstein and Prelec's (1992) Sub-Proportional Utility Function and Magnitude-Independent Discount Factor

The preference is represented by the intertemporal utility function

$$U(x,y) = u(x) + \delta u(y)$$

where u is sub-proportional, i.e.

$$x_1 < x_2 \text{ and } k > 1 \Rightarrow \frac{u(x_1)}{u(x_2)}\frac{u(kx_2)}{u(kx_1)} > 1.$$

The GDF is still

$$\delta(y) = \delta.$$

The EIS is

$$\varepsilon(x,y) = \left[-\frac{u''(y)y}{u'(y)}\frac{u'(x)x}{u'(x)x + \delta u'(y)y} - \frac{u''(x)x}{u'(x)}\frac{\delta u'(y)y}{u'(x)x + \delta u'(y)y}\right]^{-1}.$$

The relation between the EIS and the stake depends on the functional form of $u(\cdot)$. For instance, if the period utility function is a power utility function with positive background consumption, as in Appendix 3.A.2 (it is easy to show its sub-proportionality), the EIS is decreasing in the stake. If, alternatively, the period utility function is

$$u(x) = x + \frac{x^\alpha}{\alpha}$$

where $\alpha < 1$, the sub-proportionality can be seen from the fact that

$$\frac{u(x)}{u(kx)} = \frac{x + \frac{x^\alpha}{\alpha}}{kx + \frac{(kx)^\alpha}{\alpha}}$$

is decreasing in x. In this case, the EIS is

$$\varepsilon(x,y) = \left[-\frac{u''(y)y}{u'(y)}\frac{u'(x)x}{u'(x)x + \delta u'(y)y} - \frac{u''(x)x}{u'(x)}\frac{\delta u'(y)y}{u'(x)x + \delta u'(y)y}\right]^{-1}.$$

Since

$$-\frac{u''(x)x}{u'(x)} = (1-\alpha)\frac{x^{\alpha-1}}{1 + x^{\alpha-1}}$$

is decreasing in x, the EIS is increasing in the stake.

3.A.5 A Mental Accounting Version of Fudenberg and Levine's (2006) Bank-Nightclub Model

As introduced in Section 3.3.3, the model is a classic discounting model with the following period utility function:

$$u(z) = \begin{cases} \dfrac{(z+\omega)^\alpha}{\alpha}, & \text{if } z \le \left[(1+\gamma)^{\frac{1}{1-\alpha}} - 1\right]\omega \\ \left(\dfrac{\delta}{1-\delta} + (1+\gamma)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \dfrac{\left(z + \frac{\omega}{1-\delta}\right)^\alpha}{\alpha} - \gamma\dfrac{(z+\omega)^\alpha}{\alpha} - \dfrac{\omega^\alpha}{(1-\delta)\alpha}, & \text{if } z > \left[(1+\gamma)^{\frac{1}{1-\alpha}} - 1\right]\omega \end{cases}$$

where $\gamma \ge 0$, $\alpha < 1$ and $\omega > 0$.

Since it is a classic discounting model, the GDF is a constant, δ.

Regarding the EIS, first notice that because the utility function is separated into two cases, the EIS is separated into four cases, depending if x and y are above or below the cutoff, respectively. For simplicity, we only consider the EIS at balanced profiles:

$$\varepsilon(z,z) = -\frac{u'(z)}{u''(z)z} = \begin{cases} \dfrac{1}{1-\alpha}\dfrac{z+\omega}{z}, & \text{if } z \le \left[(1+\gamma)^{\frac{1}{1-\alpha}} - 1\right]\omega \\ \dfrac{1}{1-\alpha}\dfrac{\left(\frac{\delta}{1-\delta} + (1+\gamma)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}\left(z+\frac{\omega}{1-\delta}\right)^{\alpha-1} - \gamma(z+\omega)^{\alpha-1}}{\left(\frac{\delta}{1-\delta} + (1+\gamma)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}\left(z+\frac{\omega}{1-\delta}\right)^{\alpha-2} z - \gamma(z+\omega)^{\alpha-2} z}, & \text{if } z > \left[(1+\gamma)^{\frac{1}{1-\alpha}} - 1\right]\omega \end{cases}.$$

It can be seen that the EIS in the first case is the same as the classic discounting model with a power utility function and a positive background consumption, which is decreasing in the stake and converging to $\frac{1}{1-\alpha}$. In the second case, the EIS is decreasing in the stake, and converging to $\frac{1}{1-\alpha}$, too. At the cutoff, the EIS in the first case is

$$\frac{1}{1-\alpha} \cdot \frac{(1+\gamma)^{\frac{1}{1-\alpha}}}{(1+\gamma)^{\frac{1}{1-\alpha}}-1},$$

while in the second case it is

$$\frac{1}{1-\alpha} \cdot \frac{1-\frac{\gamma}{1+\gamma}}{\frac{(1+\gamma)^{\frac{1}{1-\alpha}}-1}{\frac{\delta}{1-\delta}+(1+\gamma)^{\frac{1}{1-\alpha}}}-\frac{\gamma}{1+\gamma}\frac{(1+\gamma)^{\frac{1}{1-\alpha}}-1}{(1+\gamma)^{\frac{1}{1-\alpha}}}} > \frac{1}{1-\alpha} \cdot \frac{(1+\gamma)^{\frac{1}{1-\alpha}}}{(1+\gamma)^{\frac{1}{1-\alpha}}-1}.$$

Hence, there is an upward jump in the EIS at the cutoff.

3.A.6 *Parametric Form of Noor's (2011) Magnitude-Dependent Discounting Model*

The preference is represented by the utility function

$$U(x,y) = d^{\frac{t_1}{x^\beta}} \frac{1}{\alpha} x^\alpha + d^{\frac{t_2}{y^\beta}} \frac{1}{\alpha} y^\alpha$$

where $\beta < \alpha < 1$ and $d < 1$. The function is additively separable. The first-order derivative w.r.t. y is

$$U_2(x,y) = d^{\frac{t_2}{y^\beta}} y^{\alpha-1} - \frac{\beta t_2 \ln d}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-\beta-1}.$$

The second-order derivative w.r.t. y is

$$U_{22}(x,y) = d^{\frac{t_2}{y^\beta}} (\alpha-1) y^{\alpha-2} - (2\alpha-\beta-1)\frac{\beta t_2 \ln d}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-\beta-2} + \frac{(\beta t_2 \ln d)^2}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-2\beta-2}.$$

Therefore, it holds that

$$-\frac{U_{22}y}{U_2} = \frac{(1-\alpha)d^{\frac{t_2}{y^\beta}} y^{\alpha-1} + (2\alpha-\beta-1)\frac{\beta t_2 \ln d}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-\beta-1} - \frac{(\beta t_2 \ln d)^2}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-2\beta-1}}{d^{\frac{t_2}{y^\beta}} y^{\alpha-1} - \frac{\beta t_2 \ln d}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-\beta-1}}$$

$$= \frac{(1-\alpha) + (2\alpha - \beta - 1)\frac{\beta t_2 \ln d}{\alpha} y^{-\beta} - \frac{(\beta t_2 \ln d)^2}{\alpha} y^{-2\beta}}{1 - \frac{\beta t_2 \ln d}{\alpha} y^{-\beta}}.$$

The denominator is decreasing in y, while the numerator is increasing in y. Hence, the ratio is increasing in y. The limits are

$$-\lim_{y \to 0} \frac{U_{22} y}{U_2} = -\infty$$

and

$$-\lim_{y \to \infty} \frac{U_{22} y}{U_2} = 1 - \alpha.$$

Similarly, $-\frac{U_{11} x}{U_1}$ is increasing in x, and changing from $-\infty$ to $1-\alpha$.

The GDF is

$$\delta(y) = \frac{d^{\frac{t_2}{y^\beta}} y^{\alpha-1} - \frac{\beta t_2 \ln d}{\alpha} d^{\frac{t_2}{y^\beta}} y^{\alpha-\beta-1}}{d^{\frac{t_1}{y^\beta}} y^{\alpha-1} - \frac{\beta t_1 \ln d}{\alpha} d^{\frac{t_1}{y^\beta}} y^{\alpha-\beta-1}}$$

which is increasing in y and changes from 0 to 1.

The EIS is

$$\varepsilon(x, y) = \left[-\frac{U_{22}(y) y}{U_2(y)} \cdot \frac{U_1(x) x}{U_1(x) x + U_2(y) y} - \frac{U_{22}(x) x}{U_2(x)} \cdot \frac{U_2(y) y}{U_1(x) x + U_2(y) y} \right]^{-1}.$$

Since

$$\frac{U_1(x) x}{U_1(x) x + U_2(y) y} + \frac{U_2(y) y}{U_1(x) x + U_2(y) y} = 1,$$

the EIS decreases from 0 to $-\infty$, and then decreases from $+\infty$ to 0. It means that, as the stake increases from 0 to ∞, the indifference curve changes from infinitely concave to mildly convex with an EIS of $\frac{1}{1-\alpha}$.

3.A.7 Benhabib, Bisin and Schotter's (2010) Fixed-cost-of-waiting Model

The preference is represented by a utility function

$$U(x,y)=\begin{cases}\dfrac{1}{\alpha}x^{\alpha}, & \text{if } y=0\\ \dfrac{1}{\alpha}x^{\alpha}+\delta\dfrac{1}{\alpha}y^{\alpha}-C, & \text{if } y>0\end{cases}$$

where $C>0$ is a fixed cost.

The model is the same as the classic discounting model with a power utility function. Therefore, neither the GDF nor the EIS changes with the stake. The only difference is that this model predicts the all-sooner effect. To see this, we derive the ARS between a balanced profile, (z,z), and a pure sooner reward, $(x,0)$, with the same utility. Given the same utility, it holds that

$$\frac{1}{\alpha}z^{\alpha}+\delta\frac{1}{\alpha}z^{\alpha}-C=\frac{1}{\alpha}x^{\alpha},$$

which implies

$$x=[(1+\delta)z^{\alpha}-\alpha C]^{\frac{1}{\alpha}}.$$

Then the ARS between the two profiles is

$$\begin{aligned}ARS(z,z,\infty)&=-\frac{z-x}{z}\\&=\left[(1+\delta)-\frac{\alpha C}{z^{\alpha}}\right]^{\frac{1}{\alpha}}-1.\end{aligned}$$

As a benchmark, the ARS between a balanced profile and a mixed profile very close to the sooner boundary is

$$\lim_{\lambda\to\infty}ARS(z,z,\lambda)=[(1+\delta)]^{\frac{1}{\alpha}}-1.$$

Apparently, the ARS to the sooner boundary is less than that to a mixed profile near the boundary, and hence the all-sooner effect exists.

When the stake increases (i.e. z increases), the difference between the two ARSs is decreasing and converges to 0, which implies that the all-sooner effect is less pronounced for a higher stake.

3.A.8 A Simple Version of Holden and Quiggin's (2017) Zooming Model

Holden and Quiggin (2017) propose the zooming model. The main idea is that the background consumption is increasing in the stake. The general model is

$$U(x,y) = u\big(x + f(x,y)\big) + \delta u\big(y + f(x,y)\big).$$

In their study, they assumed the background consumption is a function of the pure later reward in a binary choice problem, i.e. $f(x,y) = g(y)$. We adapt their model to binary choices between two-outcome profiles, and hence we assume the background consumption is a function of the total amount of a profile. In the meantime, we assume a simple period utility function, and hence the preference is represented by

$$U(x,y) = \frac{\big(x + b(x+y)^{\beta}\big)^{\alpha}}{\alpha} + \delta \frac{\big(y + b(x+y)^{\beta}\big)^{\alpha}}{\alpha}$$

where $\beta < \alpha < 1$.

We derive the first and second derivatives:

$$U_1 = (x + B_0)^{\alpha-1}(1 + B_1) + \delta(y + B_0)^{\alpha-1}B_1$$

$$U_2 = (x + B_0)^{\alpha-1}B_1 + \delta(y + B_0)^{\alpha-1}(1 + B_1)$$

$$\begin{aligned} U_{11} = {} & (x + B_0)^{\alpha-2}(\alpha - 1)(1 + B_1)^2 + (x + B_0)^{\alpha-1}B_2 \\ & + \delta(y + B_0)^{\alpha-2}(\alpha - 1)B_1^2 + \delta(y + B_0)^{\alpha-1}B_2 \end{aligned}$$

$$\begin{aligned} U_{12} = {} & (x + B_0)^{\alpha-2}(\alpha - 1)B_1(1 + B_1) + (x + B_0)^{\alpha-1}B_2 \\ & + \delta(y + B_0)^{\alpha-2}(\alpha - 1)(1 + B_1)B_1 + \delta(y + B_0)^{\alpha-1}B_2 \end{aligned}$$

$$\begin{aligned} U_{22} = {} & (x + B_0)^{\alpha-2}(\alpha - 1)B_1^2 + (x + B_0)^{\alpha-1}B_2 \\ & + \delta(y + B_0)^{\alpha-2}(\alpha - 1)(1 + B_1)^2 + \delta(y + B_0)^{\alpha-1}B_2 \end{aligned}$$

where $B_0 = b(x+y)^\beta$, $B_1 = b\beta(x+y)^{\beta-1}$, and $B_2 = b\beta(\beta-1)(x+y)^{\beta-2}$.

The GDF is

$$\delta(y) = \frac{\delta + B_1 + \delta B_1}{1 + B_1 + \delta B_1},$$

which is decreasing in y and changes from 1 to δ.

The EIS is

$$\begin{aligned}\varepsilon(x,y) &= \frac{U_1'U_2'[U_1'x + U_2'y]}{[U_{21}''U_1'U_2' - U_{22}''(U_1')^2 - U_{11}''(U_2')^2 + U_{12}''U_1'U_2']xy} \\ &= \left[\frac{2U_{12}''xy}{U_1'x + U_2'y} - \frac{U_{22}''U_1'xy}{U_2'[U_1'x + U_2'y]} - \frac{U_{11}''U_2'xy}{U_1'[U_1'x + U_2'y]}\right]^{-1}.\end{aligned}$$

It is hard to see the monotonicity analytically. We thereby investigate it numerically. By experimenting with the parameters in the following ranges: $\beta \in [0.001,0.999], \alpha \in [0.001,0.999], \delta \in [0.001,0.999], b \in (0,100000], \frac{x}{y} \in [0.001,1000]$, we find that the EIS is decreasing in the stake and converges to $\frac{1}{1-\alpha}$. □

3.B Proofs of the Predictions of the Simple Model

In Section 3.5.3, we provide a model represented by the following utility function:

$$U(x,y) = \begin{cases} (1-\mu)x + \mu\frac{(x+\omega)^\alpha}{\alpha} + \delta_0\mu\frac{\omega^\alpha}{\alpha}, & \text{if } y = 0 \\ (1-\mu)x + \mu\frac{(x+\omega)^\alpha}{\alpha} + \delta_\infty(1-\mu)y + \delta_0\mu\frac{(y+\omega)^\alpha}{\alpha} - C, & \text{if } y > 0 \end{cases}$$

where $\mu \in [0,1], 0 \le \delta_0 \le \delta_\infty \le 1, \alpha < 1, C > 0$.

In this section, we show that the model has the following characteristics: (i) the indifference curves are convex in the interior; (ii) the all-sooner effect exists, and its effect on the ARS is decreasing in the stake; (iii) the EIS first decreases and then increases to ∞ as the stake increases from 0 to ∞; (iv) the GDF changes monotonically from δ_0 to δ_∞ as the stake increases from 0 to ∞.

(i): Notice that the utility function is additively separable. Thereby, if each period utility function is concave in the interior, the indifference curves are convex in the interior. The second derivative of the first period utility function is

$$-\mu(1-\alpha)(x+\omega)^{\alpha-2} < 0.$$

The second derivative of the second period utility function is

$$-\delta_0\mu(1-\alpha)(y+\omega)^{\alpha-2} < 0.$$

The convexity in the interior is thus proved.

(ii): Since there is a fixed cost, the proof is the same as in Appendix 3.A.7. The all-sooner effect exists and its effect on the ARS is decreasing in the stake.

(iii): For the first period utility function, it holds that

$$-\frac{U_{11}x}{U_1} = (1-\alpha)\frac{\mu(x+\omega)^{\alpha-2}x}{1-\mu+\mu(x+\omega)^{\alpha-1}}.$$

It is first increasing and then decreasing in x and converges to 0. Similarly, for the second period utility function, it holds that $-\frac{U_{22}y}{U_2}$ is first increasing and then decreasing in y and converges to 0. The EIS is thereby first decreasing and then increasing in the stake and converges to ∞. It means that the indifference curve is initially the same as that of a classic discounting model with power utility functions and positive background consumption, but then it becomes less concave as the stake becomes higher, and it converges to a straight line.

(iv): The GDF is

$$\delta(y) = \frac{(1-\mu)\delta_\infty + \mu\delta_0(y+\omega)^{\alpha-1}}{1-\mu+\mu(y+\omega)^{\alpha-1}},$$

which is non-decreasing in y.

Hence,

$$\lim_{y\to 0}\delta(y) = \frac{(1-\mu)\delta_\infty + \mu\delta_0\omega^{\alpha-1}}{1-\mu+\mu\omega^{\alpha-1}}$$

which is a weighted average of δ_∞ and δ_0, and

$$\lim_{y \to \infty} \delta(y) = \delta_{\infty}. \square$$

3.C Parameters Used in the Choice Lists

Choice list 1: to the later boundary, for the lower stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	0.00	60.00	20.00	20.00	
2	0.00	56.80	20.00	20.00	1.92
3	0.00	53.80	20.00	20.00	1.76
4	0.00	52.40	20.00	20.00	1.66
5	0.00	51.10	20.00	20.00	1.59
6	0.00	49.80	20.00	20.00	1.52
7	0.00	48.60	20.00	20.00	1.46
8	0.00	47.40	20.00	20.00	1.40
9	0.00	46.30	20.00	20.00	1.34
10	0.00	45.20	20.00	20.00	1.29
11	0.00	44.20	20.00	20.00	1.23
12	0.00	43.20	20.00	20.00	1.18
13	0.00	42.20	20.00	20.00	1.13
14	0.00	41.30	20.00	20.00	1.09
15	0.00	40.40	20.00	20.00	1.04
16	0.00	39.60	20.00	20.00	1.00
17	0.00	38.80	20.00	20.00	0.96
18	0.00	38.00	20.00	20.00	0.92
19	0.00	37.30	20.00	20.00	0.88
20	0.00	36.60	20.00	20.00	0.85
21	0.00	35.90	20.00	20.00	0.81
22	0.00	34.60	20.00	20.00	0.76
23	0.00	34.00	20.00	20.00	0.71
24	0.00	33.40	20.00	20.00	0.69
25	0.00	32.30	20.00	20.00	0.64
26	0.00	31.80	20.00	20.00	0.60
27	0.00	30.40	20.00	20.00	0.56
28	0.00	29.20	20.00	20.00	0.49
29	0.00	24.00	20.00	20.00	0.30
30	0.00	20.00	20.00	20.00	

Choice list 2: to the interior in the future-leaning domain, for the lower stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	15.00	30.00	20.00	20.00	
2	14.80	29.60	20.00	20.00	1.92
3	14.60	29.20	20.00	20.00	1.76
4	14.50	29.00	20.00	20.00	1.66
5	14.40	28.80	20.00	20.00	1.59
6	14.30	28.60	20.00	20.00	1.52
7	14.20	28.40	20.00	20.00	1.46
8	14.10	28.20	20.00	20.00	1.40
9	14.00	28.00	20.00	20.00	1.34
10	13.90	27.80	20.00	20.00	1.29
11	13.80	27.60	20.00	20.00	1.23
12	13.70	27.40	20.00	20.00	1.18
13	13.60	27.20	20.00	20.00	1.13
14	13.50	27.00	20.00	20.00	1.09
15	13.40	26.80	20.00	20.00	1.04
16	13.30	26.60	20.00	20.00	1.00
17	13.20	26.40	20.00	20.00	0.96
18	13.10	26.20	20.00	20.00	0.92
19	13.00	26.00	20.00	20.00	0.88
20	12.90	25.80	20.00	20.00	0.85
21	12.80	25.60	20.00	20.00	0.81
22	12.70	25.40	20.00	20.00	0.76
23	12.60	25.20	20.00	20.00	0.71
24	12.50	25.00	20.00	20.00	0.69
25	12.40	24.80	20.00	20.00	0.64
26	12.30	24.60	20.00	20.00	0.60
27	12.10	24.20	20.00	20.00	0.56
28	11.90	23.80	20.00	20.00	0.49
29	10.90	21.80	20.00	20.00	0.30
30	10.00	20.00	20.00	20.00	

Choice list 3: to the interior in the present-leaning domain, for the lower stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	30.00	15.00	20.00	20.00	
2	29.60	14.80	20.00	20.00	0.52
3	29.20	14.60	20.00	20.00	0.57
4	29.00	14.50	20.00	20.00	0.60
5	28.80	14.40	20.00	20.00	0.63
6	28.60	14.30	20.00	20.00	0.66
7	28.40	14.20	20.00	20.00	0.69
8	28.20	14.10	20.00	20.00	0.71
9	28.00	14.00	20.00	20.00	0.75
10	27.80	13.90	20.00	20.00	0.78
11	27.60	13.80	20.00	20.00	0.81
12	27.40	13.70	20.00	20.00	0.85
13	27.20	13.60	20.00	20.00	0.88
14	27.00	13.50	20.00	20.00	0.92
15	26.80	13.40	20.00	20.00	0.96
16	26.60	13.30	20.00	20.00	1.00
17	26.40	13.20	20.00	20.00	1.04
18	26.20	13.10	20.00	20.00	1.09
19	26.00	13.00	20.00	20.00	1.13
20	25.80	12.90	20.00	20.00	1.18
21	25.60	12.80	20.00	20.00	1.23
22	25.40	12.70	20.00	20.00	1.31
23	25.20	12.60	20.00	20.00	1.40
24	25.00	12.50	20.00	20.00	1.46
25	24.80	12.40	20.00	20.00	1.55
26	24.60	12.30	20.00	20.00	1.66
27	24.20	12.10	20.00	20.00	1.80
28	23.80	11.90	20.00	20.00	2.04
29	21.80	10.90	20.00	20.00	3.31
30	20.00	10.00	20.00	20.00	

Choice list 4: to the sooner boundary, for the lower stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	60.00	0.00	20.00	20.00	
2	56.80	0.00	20.00	20.00	0.52
3	53.80	0.00	20.00	20.00	0.57
4	52.40	0.00	20.00	20.00	0.60
5	51.10	0.00	20.00	20.00	0.63
6	49.80	0.00	20.00	20.00	0.66
7	48.60	0.00	20.00	20.00	0.69
8	47.40	0.00	20.00	20.00	0.71
9	46.30	0.00	20.00	20.00	0.75
10	45.20	0.00	20.00	20.00	0.78
11	44.20	0.00	20.00	20.00	0.81
12	43.20	0.00	20.00	20.00	0.85
13	42.20	0.00	20.00	20.00	0.88
14	41.30	0.00	20.00	20.00	0.92
15	40.40	0.00	20.00	20.00	0.96
16	39.60	0.00	20.00	20.00	1.00
17	38.80	0.00	20.00	20.00	1.04
18	38.00	0.00	20.00	20.00	1.09
19	37.30	0.00	20.00	20.00	1.13
20	36.60	0.00	20.00	20.00	1.18
21	35.90	0.00	20.00	20.00	1.23
22	34.60	0.00	20.00	20.00	1.31
23	34.00	0.00	20.00	20.00	1.40
24	33.40	0.00	20.00	20.00	1.46
25	32.30	0.00	20.00	20.00	1.55
26	31.80	0.00	20.00	20.00	1.66
27	30.40	0.00	20.00	20.00	1.80
28	29.20	0.00	20.00	20.00	2.04
29	24.00	0.00	20.00	20.00	3.31
30	20.00	0.00	20.00	20.00	

Choice list 5: to the later boundary, for the higher stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	0.00	240.00	80.00	80.00	
2	0.00	227.20	80.00	80.00	1.92
3	0.00	215.20	80.00	80.00	1.76
4	0.00	209.60	80.00	80.00	1.66
5	0.00	204.40	80.00	80.00	1.59
6	0.00	199.20	80.00	80.00	1.52
7	0.00	194.40	80.00	80.00	1.46
8	0.00	189.60	80.00	80.00	1.40
9	0.00	185.20	80.00	80.00	1.34
10	0.00	180.80	80.00	80.00	1.29
11	0.00	176.80	80.00	80.00	1.23
12	0.00	172.80	80.00	80.00	1.18
13	0.00	168.80	80.00	80.00	1.13
14	0.00	165.20	80.00	80.00	1.09
15	0.00	161.60	80.00	80.00	1.04
16	0.00	158.40	80.00	80.00	1.00
17	0.00	155.20	80.00	80.00	0.96
18	0.00	152.00	80.00	80.00	0.92
19	0.00	149.20	80.00	80.00	0.88
20	0.00	146.40	80.00	80.00	0.85
21	0.00	143.60	80.00	80.00	0.81
22	0.00	138.40	80.00	80.00	0.76
23	0.00	136.00	80.00	80.00	0.71
24	0.00	133.60	80.00	80.00	0.69
25	0.00	129.20	80.00	80.00	0.64
26	0.00	127.20	80.00	80.00	0.60
27	0.00	121.60	80.00	80.00	0.56
28	0.00	116.80	80.00	80.00	0.49
29	0.00	96.00	80.00	80.00	0.30
30	0.00	80.00	80.00	80.00	

Choice list 6: to the interior in the future-leaning domain, for the higher stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	60.00	120.00	80.00	80.00	
2	59.20	118.40	80.00	80.00	1.92
3	58.40	116.80	80.00	80.00	1.76
4	58.00	116.00	80.00	80.00	1.66
5	57.60	115.20	80.00	80.00	1.59
6	57.20	114.40	80.00	80.00	1.52
7	56.80	113.60	80.00	80.00	1.46
8	56.40	112.80	80.00	80.00	1.40
9	56.00	112.00	80.00	80.00	1.34
10	55.60	111.20	80.00	80.00	1.29
11	55.20	110.40	80.00	80.00	1.23
12	54.80	109.60	80.00	80.00	1.18
13	54.40	108.80	80.00	80.00	1.13
14	54.00	108.00	80.00	80.00	1.09
15	53.60	107.20	80.00	80.00	1.04
16	53.20	106.40	80.00	80.00	1.00
17	52.80	105.60	80.00	80.00	0.96
18	52.40	104.80	80.00	80.00	0.92
19	52.00	104.00	80.00	80.00	0.88
20	51.60	103.20	80.00	80.00	0.85
21	51.20	102.40	80.00	80.00	0.81
22	50.80	101.60	80.00	80.00	0.76
23	50.40	100.80	80.00	80.00	0.71
24	50.00	100.00	80.00	80.00	0.69
25	49.60	99.20	80.00	80.00	0.64
26	49.20	98.40	80.00	80.00	0.60
27	48.40	96.80	80.00	80.00	0.56
28	47.60	95.20	80.00	80.00	0.49
29	43.60	87.20	80.00	80.00	0.30
30	40.00	80.00	80.00	80.00	

Choice list 7: to the interior in the present-leaning domain, for the higher stake

Problem	LEFT Tomorrow (euro)	LEFT 20 WEEKS + 1 Day (euro)	RIGHT Tomorrow (euro)	RIGHT 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	120.00	60.00	80.00	80.00	
2	118.40	59.20	80.00	80.00	0.52
3	116.80	58.40	80.00	80.00	0.57
4	116.00	58.00	80.00	80.00	0.60
5	115.20	57.60	80.00	80.00	0.63
6	114.40	57.20	80.00	80.00	0.66
7	113.60	56.80	80.00	80.00	0.69
8	112.80	56.40	80.00	80.00	0.71
9	112.00	56.00	80.00	80.00	0.75
10	111.20	55.60	80.00	80.00	0.78
11	110.40	55.20	80.00	80.00	0.81
12	109.60	54.80	80.00	80.00	0.85
13	108.80	54.40	80.00	80.00	0.88
14	108.00	54.00	80.00	80.00	0.92
15	107.20	53.60	80.00	80.00	0.96
16	106.40	53.20	80.00	80.00	1.00
17	105.60	52.80	80.00	80.00	1.04
18	104.80	52.40	80.00	80.00	1.09
19	104.00	52.00	80.00	80.00	1.13
20	103.20	51.60	80.00	80.00	1.18
21	102.40	51.20	80.00	80.00	1.23
22	101.60	50.80	80.00	80.00	1.31
23	100.80	50.40	80.00	80.00	1.40
24	100.00	50.00	80.00	80.00	1.46
25	99.20	49.60	80.00	80.00	1.55
26	98.40	49.20	80.00	80.00	1.66
27	96.80	48.40	80.00	80.00	1.80
28	95.20	47.60	80.00	80.00	2.04
29	87.20	43.60	80.00	80.00	3.31
30	80.00	40.00	80.00	80.00	

Choice list 8: to the sooner boundary, for the higher stake

Problem	LEFT option Tomorrow (euro)	LEFT option 20 WEEKS + 1 Day (euro)	RIGHT option Tomorrow (euro)	RIGHT option 20 WEEKS + 1 Day (euro)	ARS implied by a switching
1	240.00	0.00	80.00	80.00	
2	227.20	0.00	80.00	80.00	0.52
3	215.20	0.00	80.00	80.00	0.57
4	209.60	0.00	80.00	80.00	0.60
5	204.40	0.00	80.00	80.00	0.63
6	199.20	0.00	80.00	80.00	0.66
7	194.40	0.00	80.00	80.00	0.69
8	189.60	0.00	80.00	80.00	0.71
9	185.20	0.00	80.00	80.00	0.75
10	180.80	0.00	80.00	80.00	0.78
11	176.80	0.00	80.00	80.00	0.81
12	172.80	0.00	80.00	80.00	0.85
13	168.80	0.00	80.00	80.00	0.88
14	165.20	0.00	80.00	80.00	0.92
15	161.60	0.00	80.00	80.00	0.96
16	158.40	0.00	80.00	80.00	1.00
17	155.20	0.00	80.00	80.00	1.04
18	152.00	0.00	80.00	80.00	1.09
19	149.20	0.00	80.00	80.00	1.13
20	146.40	0.00	80.00	80.00	1.18
21	143.60	0.00	80.00	80.00	1.23
22	138.40	0.00	80.00	80.00	1.31
23	136.00	0.00	80.00	80.00	1.40
24	133.60	0.00	80.00	80.00	1.46
25	129.20	0.00	80.00	80.00	1.55
26	127.20	0.00	80.00	80.00	1.66
27	121.60	0.00	80.00	80.00	1.80
28	116.80	0.00	80.00	80.00	2.04
29	96.00	0.00	80.00	80.00	3.31
30	80.00	0.00	80.00	80.00	

4 Does Making a Choice Affect Subsequent Belief Formation? An Experimental Study

In economic theories, when uncertainty is present, choices are based on beliefs and preferences. While substantial progress has been achieved in understanding how people make choices given their beliefs, economists have paid less attention to the question in the opposite direction: does making a choice have an effect on subsequent belief formation?

One example would be John choosing his dessert. Does choosing yoghurt over ice cream make John believe in a larger health benefit of yoghurt relative to ice cream? This could happen either because believing that brings a higher anticipatory utility (e.g. yoghurt brings less risk of obesity in the future) or because believing that raises self-esteem (e.g. John made a correct choice).

This possibility has important implications in both positive and normative terms. If making a choice shifts the decision maker's belief more in favor of the chosen option, it may lead to persistence of early-stage beliefs and behaviors. When such persistence exists, beliefs do not necessarily converge to the truth. In the example of John, if choosing yoghurt makes John believe in a larger benefit of yoghurt, he is more likely to choose yoghurt again. Choosing yoghurt repeatedly makes John's belief continually shift in favor of yoghurt. Then even if the truth is that the health benefit of yoghurt is not large enough to overweight its disadvantage in deliciousness, his belief might converge to one that supports the choice of yoghurt. In this case, intervention programs that change behaviors in the short-run have a long-run effect on beliefs, and thus can be used to induce desired beliefs and behaviors in the long-run. Such an intervention program is justified when John has inconsistent preferences: the Cold self (i.e. the self when the temptation is absent) is more patient than the Hot self (i.e. the self when the temptation is present) but has no direct control over choices (Bernheim and Rangel 2004). If an intervention program is available, the Cold self is willing to participate in order to induce a long-run belief that supports the choice of yoghurt.

In this paper, we perform an experiment to investigate whether making a choice has a direct effect on subsequent belief formation. To be more specific, we are interested in whether making a choice, compared to not making a choice, leads to any difference in future beliefs about the attractiveness of the options, due to some cognitive biases.

The effect should not be confused with the following two information effects which often occur before or after a choice is made. First, if a decision maker is asked to make a choice, she might put more effort into collecting information, which could lead to a different belief from that if she does not need to make a choice (the information-collection effect). Second, after a choice is made, the decision maker might receive some feedback from consuming the chosen option, which brings her additional information than if she does not make a choice (the feedback effect). Those two effects could be explained by Bayesian learning, and hence are not our focus. We rule them out in our experimental design.

Our study is related to but different from two strands of literature. First, there are studies on choice-induced preference change. They find evidence that the act of choosing a good (Brehm 1956), waiting for a good (Aronson and Carlsmith 1963), participating in an activity (Festinger and Carlsmith 1959; Aronson and Mills 1959; Gerard and Matthewson 1966), writing an essay to support a group of people (Cohen 1962), or voting for a candidate (Mullainathan and Washington 2009) leads to a more favorable opinion of the corresponding good, activity or person in the future (see Cooper 2007 for a review). *Preferences over alternatives* can be divided into *beliefs about attributes of alternatives* and *preferences over lotteries of attributes*. For instance, preferences over ice cream and yoghurt can be divided into beliefs about deliciousness and health consequences of the two options, and preferences over lotteries of deliciousness and health consequences. In this sense, our study can be seen as an investigation into the effect on the first part when the second part is controlled. Moreover, because preferences over goods, activities, groups of people or candidates are not verifiable, it is not possible to incentivize reported preferences directly. In contrast, we elicit beliefs about verifiable objects, and hence by incentivizing belief elicitation, we can examine whether making a choice has economically meaningful effects on future beliefs.

Second, a few studies investigate whether having a stake in some states of nature induces people to distort their beliefs. Studies on beliefs about self-relevant events, such as one's own cognitive ability or beauty, usually find evidence of distortion (Eil and Rao 2011, Ertac 2011 and Möbius et al. 2014), but studies on beliefs about objective events, such as color of a ball or weather, find mixed evidence (Ertac 2011, Barron 2018, Heger and Papageorge 2018 and the medium- and the high-incentive treatments of Coutts 2015 find no distortion, but Mayraz 2014 and the low-incentive treatment of Coutts 2015 find evidence of distortion). In those studies, what states of nature bring higher payoffs is exogenous to subjects' beliefs. In contrast, we are interested in whether beliefs are distorted when people voluntarily bet on states of nature. Whether one's bet is correct is somehow self-relevant, and hence beliefs are a priori more likely to be distorted in this case. This has potential implication to policies: an intervention program that induces people to choose a target behavior by their own might be more effective than one that forces people to perform the target behavior.

Our experiment consists of two stages. In Stage 1, subjects are presented with noisy signals about the values of two options, and then they are randomly assigned to one of the three treatments: making a choice between the two options, receiving a random option, or possessing no option. In Stage 2, subjects are presented with more signals and are asked to estimate the values of the two options. We examine whether making a choice in Stage 1 leads to a systematic change in reported beliefs in Stage 2.

We consider choice-induced belief change as a result of interaction between different psychological motives. People, on the one hand, want to form an accurate belief about values of options for potential uses in the future, and on the other hand, distort beliefs in order to derive a higher anticipatory utility or to raise self-esteem. Thereby, we expect that beliefs are more likely to be distorted when the incentive for accuracy is lower. Back to the example of John, he is less likely to admit the health risk of ice cream if it is a special hand-made product which will no longer appear in his life, because the information about health consequences will not be used any more, and hence he would rather think in a comfortable way. In our experiment, in some of the sessions, belief elicitation is incentivized while in the other sessions it is not. By comparing the results between those sessions, we can examine how choice-induced belief change is affected by the relative importance of accuracy.

We consider two potential channels of choice-induced belief change. Distortion can take place either in the retrieval of pre-choice signals (e.g. John's record of diet and weight), or in the processing of post-choice signals (e.g. magazine articles about over-eating). We thus vary the accessibility of pre-choice signals in the post-choice stage in our experiment to examine whether belief change can take place when the retrieval of pre-choice signals cannot be distorted.

Our design has three advantages. First, it rules out the selection bias caused by the inherent correlation between choices and beliefs, a problem that was not well dealt with in some previous studies (see Chen and Risen 2010 and Alós-Ferrer and Shi 2015 for detailed introductions). Since whether a subject is invited to make a choice in Stage 1 is exogenous, the study provides a clean test of choice-induced belief change. Second, belief elicitation can be incentivized in our design, which allows us to examine whether making a choice has an effect on belief formation when reported beliefs are consequential. Third, we vary the accessibility of pre-choice signals in the post-choice stage. As a result, we can test if belief change is possible when people have perfect accessibility to their pre-choice signals.

We find no significant differences between the distributions of estimates in the three groups, irrespective of whether belief elicitation is incentivized or not, and irrespective of whether pre-choice signals are perfectly accessible or not. The results suggest that in a setting where beliefs are about a single, immediately verifiable attribute, making a choice does not lead to an economically meaningful effect on subsequent belief formation.

The remaining part of the paper is structured as follows: In Section 4.1 we introduce the theoretical background of choice-induced belief change. In Section 4.2, we present our experimental design, testing procedure and testable hypotheses. We show our results in Section 4.3. In Section 4.4, we discuss how our results are related to the literature, and draw conclusions.

4.1 Theoretical Framework

We consider a two-stage situation. In Stage 1, a decision maker (henceforth DM) needs to choose between two options. She does not observe the exact values of the two options but can observe some signals. She tries her best to interpret the signals, with some noise due to cognitive limitations, and makes a choice based on her interpretation of the signals. In Stage 2, she needs to form a belief, for instance, in order to make another decision or to give a recommendation to others. At this time, she can observe her past choice, recall her pre-choice signals (perfectly or imperfectly) and also receive some new signals. For simplicity, we assume that the goal of Stage 2 generates a symmetric and single-mode incentive, so that, if there is no cognitive bias, forming an unbiased and accurate belief is the best to the DM.

However, the DM might have psychological motives other than maintaining an accurate belief, which could generate a bias in her belief. One such motive is wishful thinking: if the DM can get a higher payoff conditional on one option (e.g. Option A) having a higher value than the other (e.g. Option B), she will overestimate the value of Option A relative to Option B in order to obtain a larger anticipatory utility (Loewinstein 1987). Such a differential payoff could be caused by an exogenous event (e.g. she is given a random option) or by the DM's choice (e.g. she chooses an option).

Besides wishful thinking, the DM may also have motives that exist only if she has made a choice. First, she may have a motive to reduce cognitive dissonance, a tendency to deny her responsibility for possible bad consequences of a past choice, either by arguing that her choice was at least optimal conditional on her pre-choice signals or by arguing that her choice will turn out to be correct (Festinger 1957; see Cooper 2007 for a review). Second, she may have a motive of self-perception, a tendency to infer forgotten pre-choice signals from her choice (Bem 1965 and 1972). Both the motives could prompt the DM to shift her belief towards her choice.

Since belief change is a result of a tradeoff between the utility of accuracy and the anticipatory utility, the size of the effect is affected by the importance of accuracy (Brunnermeier and Parker 2005). If the post-choice belief is not consequential, the motive to maintain an accurate belief becomes weaker. In this case, we are more likely to observe a belief change.

There are two potential channels of belief change. The DM can either distort the retrieval of pre-choice signals or distort the processing of post-choice signals. If pre-choice signals are perfectly accessible in Stage 2, the channel of distorting pre-choice signals is no longer possible. In this case, any belief change must be caused by biased processing of post-choice signals.

4.2 Experimental Design and Analysis

4.2.1 Design

There are two stages in our experiment. Subjects are first given the instructions for Stage 1. After they finish Stage 1, the instructions for Stage 2 are delivered. The instructions are read aloud before either stage of the experiment.

In Stage 1, subjects are given two unknown numbers (called X and Y) and are asked to judge which number is larger. They cannot see the true values of the two numbers but can generate signals by clicking buttons (Figure 4.1). Subjects are informed that signals are independent, equally likely to be larger or smaller than the true value, and more likely to be close to than far away from the true value. They are also given a graph showing the probabilities of a signal deviating from the true value to a certain extent.

In fact, the true values of the two numbers are one of the following four pairs: (115, 120), (160, 165), (137, 142) and (182, 187). It is randomly determined whether X or Y is larger. Signals are independent and normally distributed around the true values, with a standard deviation of 30, and are censored when the deviation is above 100. Since the difference between X and Y is only five, signals are quite noisy. Consequently, subjects will face large uncertainty in the values of the two numbers, which provides excuses for their belief distortions.

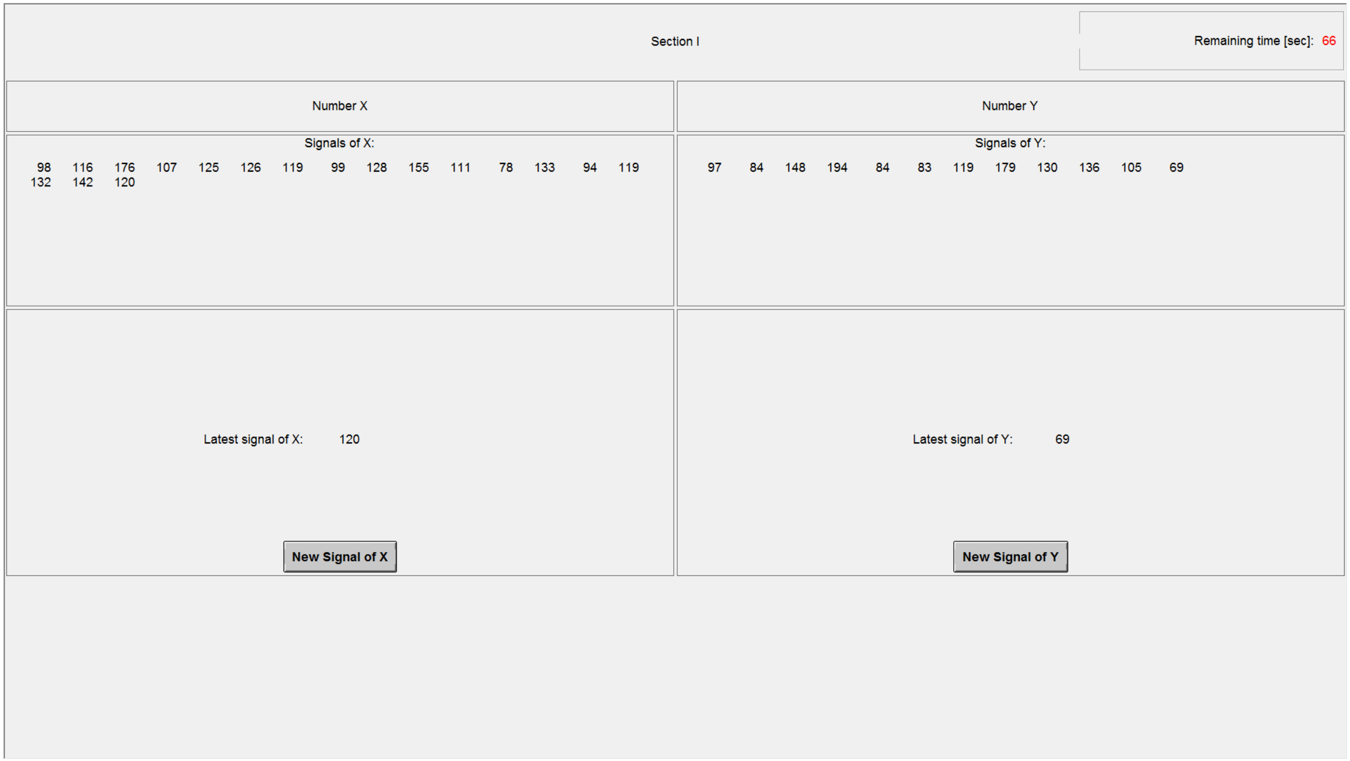


Figure 4.1: Interface of the Signal-Requesting Stage

Subjects can generate as many signals as they want within two minutes, with a minimum delay of one second between every two signals. Signals stay on the screen until time is over. After that, all signals disappear. Then subjects are randomly assigned to one of the following three treatments. In the Choice group, subjects need to choose the larger of the two numbers. If a subject makes a correct choice, she will get eight euros; otherwise she will get nothing. In the Gift group, each subject is randomly given one of the two numbers. If the given number is the larger one, she will get eight euros; otherwise she will get nothing. In the Neutral group, subjects are not assigned to any of the two numbers and each of them will get four euros. Subjects in the Choice group have 30 seconds for making a choice. If one does not make a choice in time, the computer makes a random choice for her.

When subjects generate signals, they have no idea about which treatment they will be assigned to, though they are aware of the three possibilities. In this way, the information-collection effect is ruled out.

Since the three treatments are randomly assigned, whether one needs to make a choice is exogenous to her belief. Therefore, by comparing the beliefs in the Choice group with those in the Neutral group, we can test whether there exists a choice-induced belief change.

In the Gift group, subjects are randomly assigned to X or Y. Thereby, by looking at the difference in reported beliefs between those who are assigned to the larger number and those who are assigned to the smaller number, we can test whether there is a stake-induced belief distortion in our setting.

Subjects do not know immediately after Stage 1 whether their choice or the number they are given is the larger number. In fact, they cannot see the true values of X and Y until Stage 2 is over. This implies that making a choice in Stage 1 does not bring more information about the two numbers, and as a result, the feedback effect is ruled out.

In Stage 2, subjects are asked to report their estimates of X and Y, of which the values are the same as in Stage 1. Subjects can generate signals again, but this time, for each of the two numbers, only the latest signal remains on the screen (see Figure 4.2, above the two white buttons). We restrict the accessibility to signals generated in Stage 2 in order to maximize the room for distorting beliefs.

In the middle of the screen, subjects are provided with the information about their group. Besides, subjects in the Choice group are reminded of their choices made in Stage 1, and subjects in the Gift group are reminded of the numbers assigned to them in Stage 1.

Except the information about the latest signals generated in Stage 2 and the group in Stage 1, subjects may or may not see all the signals generated in Stage 1, depending on which *condition* they are under. Under the Main condition, signals generated in Stage 1 are not displayed in Stage 2 (Figure 4.2(a)). Under the Archive condition, all signals generated in Stage 1 are displayed on the screen (Figure 4.2(b)). Given that subjects have perfect accessibility to pre-choice signals and enough time to process, the channel of distorting the retrieval of pre-choice signals is no longer possible.

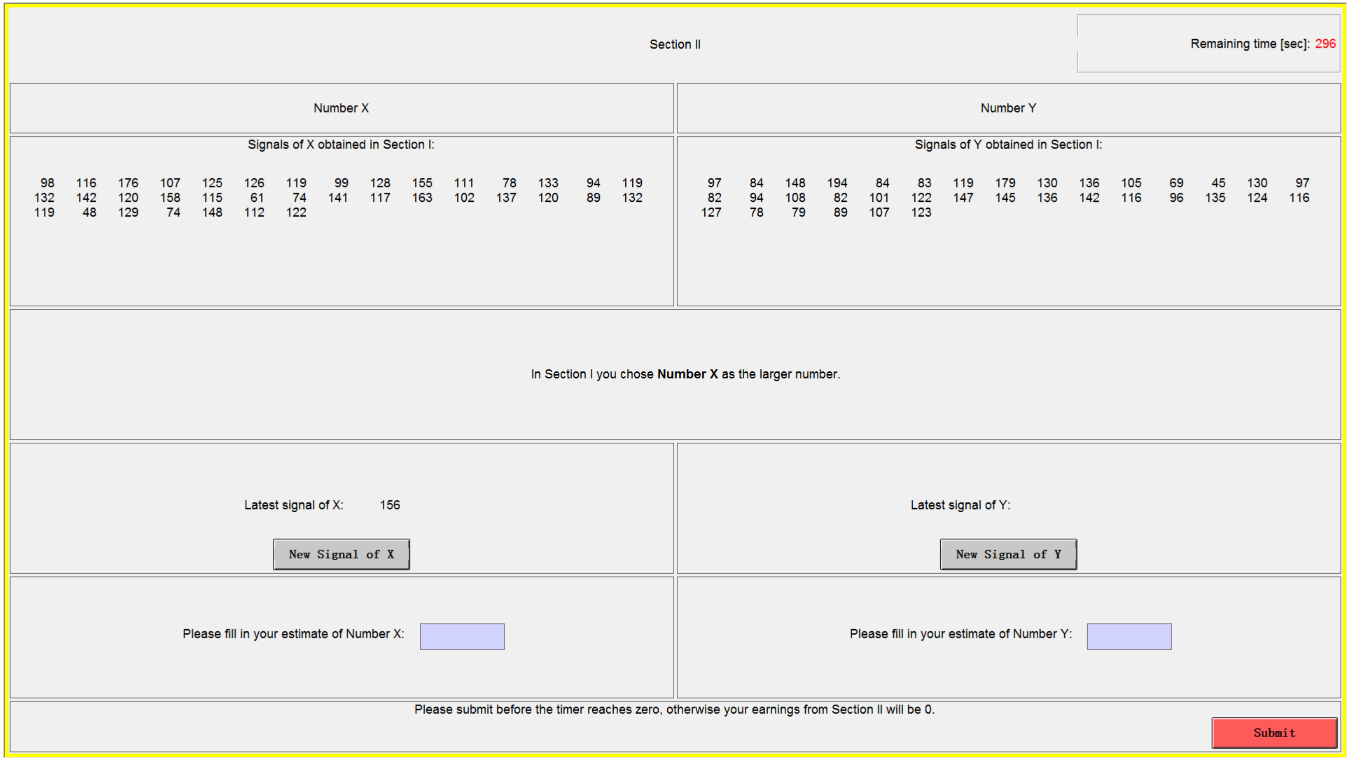
With all the information mentioned above, subjects are asked to report their estimates of both X and Y. The time limit for generating new signals and reporting beliefs is five minutes. Those who fail to submit beliefs in time get no reward from Stage 2.

Under both the Main condition and the Archive condition, we incentivize their reported beliefs with the quadratic scoring rule. The relation between one's earnings from the reported belief, m, and the absolute deviation, d, is

$$m = \max(8 - 0.02 \times d^2, 0) \text{ euros.}$$

With the incentive structure above, subjects are incentivized to report the means of their probabilistic beliefs, as long as their beliefs are symmetric.

(a) Main condition



(b) Archive condition

Figure 4.2: Interface of Stage 2

We also implement a No-Incentive condition, in which all features are the same as in the Main condition except that the reward does not depend on the reported beliefs but are fixed at four euros. It enables us to check if belief change is more likely to occur when the reported belief is not consequential.

Table 4.1 summarizes the features of the three between-subject conditions and what motives could play a role for subjects in the Choice group under the three conditions, respectively.

Table 4.1: Conditions at the Session Level and Implications to the Choice Group

	Main	Archive	No-Incentive
Incentivized?	yes	yes	no
Stage 1 signals displayed in Stage 2?	no	yes	no
Implications to the Choice group			
Maintaining accurate beliefs	Strong	Strong	Weak
Distorting the retrieval of Stage 1 signals	Possible	Not possible	Possible
Distorting the processing of Stage 2 signals	Possible	Possible	Possible

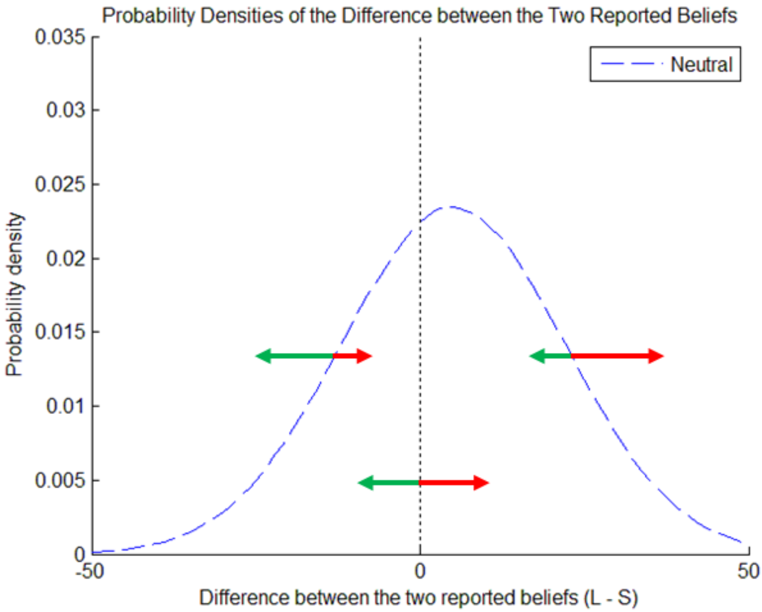
In our experiment, all subjects get three euros as a show-up fee. Besides, each subject is paid either for her earnings from Stage 1 or for her earnings from Stage 2. Which stage is paid is determined randomly at the end of the experiment. The random incentive scheme rules out the motivation of hedging.

4.2.2 Testing Procedure and Hypotheses

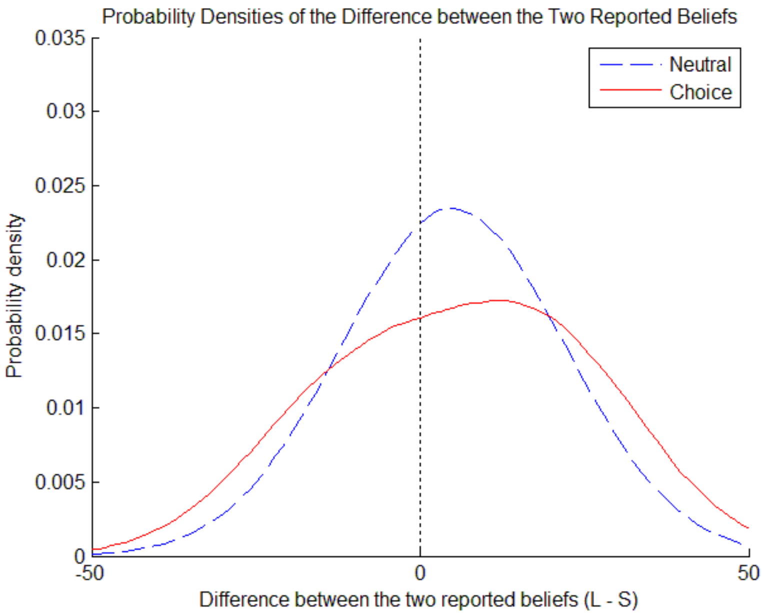
A difficulty in identifying any choice-induced effect is the inherent endogeneity in choices: subjects who believe X is larger in Stage 1 choose X, and they are more likely to believe X is larger in Stage 2 as well. Thereby a correlation between choosing X and reporting a larger estimate of X does not necessarily mean a causal effect of choices on beliefs. Chen and Risen (2010) and Alós-Ferrer and Shi (2015) show that many previous studies used testing procedures with statistical biases, either because they did not correct for the endogeneity, or they corrected in an inappropriate way.

In this paper, we develop a new approach to test the existence of a choice-induced belief change. The approach is based on the following idea: people who believe in Stage 2 that X is larger are more likely to believe in the same relation between X and Y in Stage 1 than believing in the opposite relation. Therefore, when those people are in the Choice group, they are more likely to have chosen X than having chosen Y. If making a choice makes one's belief more favorable to her choice, the people mentioned above (i.e. those who believe that X is larger in Stage 2) will on average have a belief more favorable to X, compared to their counterparts in the Neutral group. Similarly, people in the Neutral group who believe that Y is larger would have beliefs more favorable to Y were they in the Choice group. As a result, the distribution in the Choice group should be more spread out than that in the Neutral group. Figure 4.3 demonstrates the idea graphically. Here, L is one's reported belief on the truly larger number and S is her reported belief on the truly smaller number. Figure 4.3(a) shows that, for subjects with $L - \mathrm{S} > 0$, the average effect of making a choice on $L - \mathrm{S}$ is positive; for subjects with $L - \mathrm{S} < 0$, the average effect of making a choice on $L - \mathrm{S}$ is negative. Figure 4.3(b) shows the simulated distributions of beliefs on $L - \mathrm{S}$ in the Choice group and in the Neutral group: the distribution in the Choice group is more spread out than that in the Neutral group.[31]

[31] The simulation assumes the following model. Subjects receive a certain number of signals in Stage 1 and in Stage 2. Each signal is an independent and normally distributed random variable. Subjects form beliefs and make choices rationally. Making a choice shifts one's belief towards her choice by a fixed number of units. All parameters are calibrated to our sample statistics under the Main condition.



(a) Average effect of making a choice conditional on a belief



(b) Simulated distributions of beliefs in the Choice group and in the Neutral group

Figure 4.3: Simulated Effect of Making a Choice on the Distribution of Beliefs

Based on the idea above, we test the choice-induced belief change by performing a rank-sum test on the absolute belief differences over the two numbers between the Choice group and the Neutral group, i.e. we test the equality of $|L - S|$ between the Choice group and the Neutral group. If making a choice does not change one's belief, the absolute difference should be the same in the two groups. However, if a choice-induced belief change occurs, the absolute difference is supposed to be larger in the Choice group than in the Neutral group.

Based on the testing procedure above and our experimental design, we propose three hypotheses about choice-induced belief change. First, we test if making a choice leads to a belief change in a situation where the belief elicitation is incentivized and pre-choice signals are unavailable in the stage of belief formation. In this case, the channel of distorting the retrieval of pre-choice signals and that of distorting the processing of post-choice signals are both possible.

Hypothesis 1 (Choice-induced belief change when pre-choice signals are imperfectly accessible): Under the Main condition, $|L - S|$ *in the Choice group is larger than that in the Neutral group.*

We are also interested in the question whether making a choice results in a belief change when pre-choice signals are perfectly accessible. In this case, the only possible channel is to distort the processing of post-choice signals.

Hypothesis 2 (Choice-induced belief change when pre-choice signals are perfectly accessible): Under the Archive condition, $|L - S|$ *in the Choice group is larger than that in the Neutral group.*

When a reported belief is not consequential, the motive to maintain an accurate belief is weaker. We wonder if we can detect a belief change in this case, so we perform a similar test under the No-Incentive condition.

Hypothesis 3 (Choice-induced belief change when belief elicitation is not incentivized): Under the No-Incentive condition, $|L - S|$ *in the Choice group is larger than that in the Neutral group.*

To facilitate a comparison between the choice-induced belief change and the stake-induced belief distortion, we test if there is a belief distortion when a subject does not make a choice but is given a random option. For this purpose, we check if people who are randomly assigned to the larger number have a larger belief difference, $L - S$, than those who are assigned to the smaller number.[32]

Hypothesis 4 (Stake-induced belief distortion): In any condition, people who are assigned to the larger number report a larger estimate of $L - S$ *than those who are assigned to the smaller number.*

4.2.3 Implementation

Our experiment was conducted at the CentERlab, Tilburg University from March to September in 2016. 360 students of the university participated in one of the 26 sessions, 142 under the Main condition, 114 under the Archive condition and 104 under the No-Incentive condition. One session took 35 minutes on average. The average earning was €7.90.

4.3 Results

In Stage 1, we randomly assigned subjects to three treatments. 121 subjects were allocated into the Neutral Group, 118 into the Gift Group, and 121 into the Choice Group. Among those who were allocated into the Choice Group, five

[32] It is impossible to test whether choice-induced belief change is of a larger size at the individual level than stake-induced belief change by comparing L-S between the Choice group and the Gift group, because under the null hypothesis that the effects have the same size, the distribution of L-S in the Choice group spreads out from 0 while that in the Gift group spreads out from 5, which is the true value of L-S. Thereby, the medians of |L-S| in the two groups are different under the null hypothesis.

failed to make a choice within the time limit, then they were randomly given an option, and we hereby count them as in the Gift Group.

We drop a few observations which are not useful to our analysis. First, 21 subjects failed to submit their beliefs within the time limit. Second, five subjects failed to make a choice within the time limit. Third, 12 subjects reported beliefs which are out of the potential range of their signals (i.e. the difference between the reported belief and the true value is greater than 100). Those beliefs are obviously ungrounded. Fourth, two subjects requested fewer than five signals for one of the numbers in the entire experiment, which makes their reported beliefs unreliable.[33] Those four problems involve 33 subjects in total, and we hereby drop those observations.

Table 4.2 shows the number of remaining observations in each group and under each condition along with statistics about the numbers of signals requested. There are no significant differences in the number of signals requested in Stage 2 between the three groups, showing that whether one makes a choice or has a stake in an option does not affect signal-requesting behavior.

23 out of the 107 subjects (21%) in the Choice group chooses the smaller underlying number, showing that finding out the larger underlying number is not straightforward.[34] This implies that subjects face non-trivial uncertainty in the values of X and Y, which could provide excuses for subjects to distort their beliefs.

Figure 4.4 shows the comparison of absolute belief differences between the Neutral group and the Choice group. We perform rank-sum tests on those differences. Table 4.3 shows the results. Under no condition is the absolute belief difference significantly larger for the Choice group than for the Neutral group ($p>0.10$ for all conditions, one-sided). This implies that making a choice in Stage 1 does not have an effect on belief formation in Stage 2. The result holds

[33] We check the robustness of our results by including those who requested too few signals. No conclusion changes.

[34] The fraction is not very small, given the fact that the fraction would be 50% if all subjects chose randomly. The task being not straightforward is also shown by the following fact: the sample standard deviation of L-S is about 16, three times more than the true difference (which is 5).

irrespective of whether belief elicitation is incentivized or not, and irrespective of whether Stage 1 signals are perfectly accessible in Stage 2. This suggests that making a choice does not lead to any belief change in our setting, neither due to wishful thinking nor due to any choice-related motives.

Table 4.2: Number of Observations and Statistics in Each Group and under Each Condition

Condition	**Group**	**# obs**	**# signals requested in Stage 2 [std. dev.]**		**Kruskal–Wallis test of equality**
Main	Neutral	42	134.6	[52.2]	p = 0.498
	Gift	42	119.4	[62.0]	
	Choice	44	121.8	[54.3]	
Archive	Neutral	33	73.6	[40.1]	p = 0.186
	Gift	35	81.2	[36.9]	
	Choice	30	91.5	[44.0]	
No-Incentive	Neutral	36	87.0	[43.5]	p = 0.974
	Gift	32	85.7	[50.2]	
	Choice	33	85.7	[37.0]	

Notes: Number of observations, mean and standard deviation of number of signals requested in Stage 2, under each condition and in each treatment group.

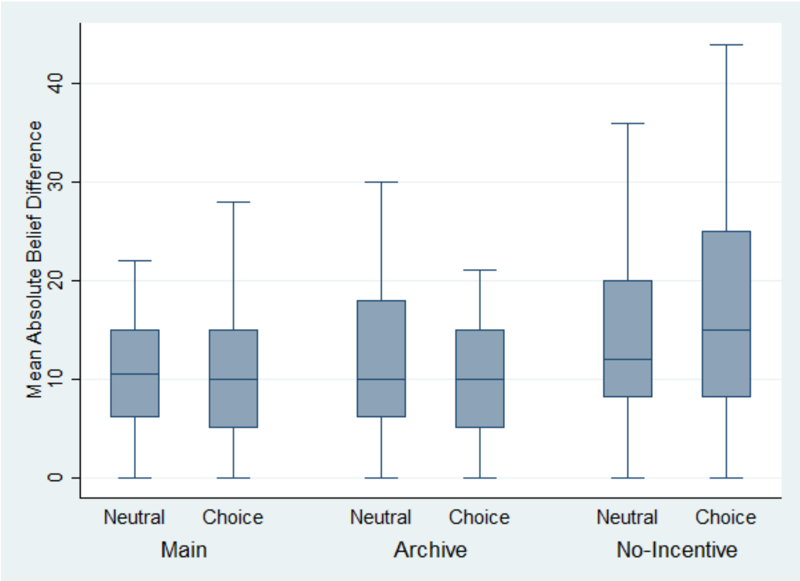


Figure 4.4: Distributions of Absolute Belief Differences in the Neutral Group and in the Choice Group

Notes: Mean absolute belief difference, $|L - S|$. For each box, the horizontal lines from top to bottom are the upper adjacent value, the 3rd quartile, the median, the 1st quartile and the lower adjacent value, respectively. Observations outside the adjacent values are dropped from the figure.

Table 4.3: Tests of Choice-Induced Belief Change

	Main	Archive	No-Incentive
Mean absolute belief difference in the Neutral Group $|D_0| = |L - S|_0$	13.14	14.15	14.94
Mean absolute belief difference in the Choice Group $|D_C| = |L - S|_C$	10.68	11.00	18.61
Rank-sum test $H_0: |D_0| = |D_C|$ $H_1: |D_0| < |D_C|$	p = 0.748	p = 0.790	p = 0.217

Notes: A mean absolute belief difference is the mean of the absolute differences between the belief on X and that on Y. This table presents the one-sided rank-sum tests on the mean absolute belief differences between the Neutral group and the Choice group under the three conditions.

Figure 4.5 shows the comparison of the belief differences between those who are given the larger number and those who are given the smaller number. We report the results of the rank-sum tests on the belief differences in Table 4.4. Under no condition are the belief differences significantly larger for those who are given the larger number as a gift than for those who are given the smaller number (p>0.10 for all conditions, two-sided).[35] Therefore, we find no evidence of stake-induce belief distortion.

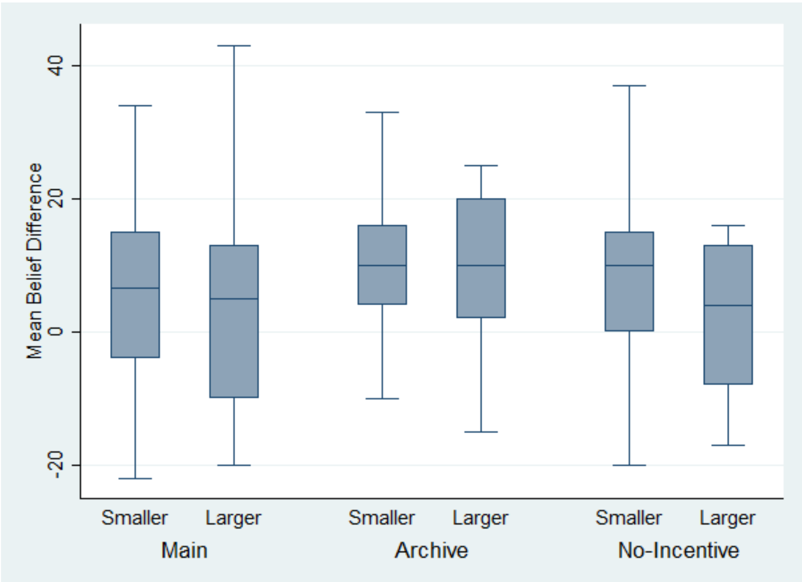


Figure 4.5: Mean Belief Differences among Those Who are Given the Larger Number and Those Who are Given the Smaller Number

Notes: Mean belief difference, $L - S$. For each box, the horizontal lines from top to bottom are the upper adjacent value, the 3rd quartile, the median, the 1st quartile and the lower adjacent value, respectively. Observations outside the adjacent values are dropped from the figure.

[35] There is, however, a trend under the Main condition and the No-Incentive condition that people who are given the larger number report a smaller belief difference than those who are given the smaller number. The reason is not likely to be hedging, not only because in theory the random incentive scheme can rule out hedging, but also because the difference is of a larger size under the No-Incentive condition where belief elicitation does not affect monetary payoffs. A possible reason is distrust about the randomness of the given option, that is, subjects may believe that they are more likely to receive the smaller number. This effect is much weaker under the Archive condition where all pre-choice signals are available again.

Table 4.4: Tests of Stake-Induced Belief Distortion

	Main	Archive	No-Incentive
Mean belief difference among those who receive the larger number as a gift $D_{GL} = (L-S)_{GL}$	4.36	11.57	-1.07
Mean belief difference among those who receive the smaller number as a gift $D_{GS} = (L-S)_{GS}$	9.20	10.36	10.71
Rank-sum test $H_0: D_{GL} = D_{GS}$ $H_1: D_{GL} \neq D_{GS}$	p = 0.203	p = 0.761	p = 0.130

Notes: A mean belief difference is the mean of the differences between the belief on the truly larger number and that on the truly smaller number. This table presents the two-sided rank-sum tests on the mean belief differences between subjects who receive the truly larger number and those who receive the truly smaller number under the three conditions.

In order to determine whether the null results are due to lack of power, we calculate the ex-ante power of our tests on choice-induced belief changes. To do so, we assume that a subject perceives signals received in the two stages as two independent and normally distributed random variables, of which the precisions are proportional to the number of signals requested in the corresponding stage. We then calibrate the parameters with the aggregate characteristics of our data. Specifically, we set the number of signals in each stage in the power analysis to be the median number of signals in the corresponding stage in our data, and set the total precision of signals to be the reciprocal of the sample variance of the belief difference in the Neutral Group.

Table 4.5 shows the results of the power analysis. If there is an effect which leads to a loss equivalent to 9% of the maximum reward in belief elicitation, our test has a chance of 85-88% at the 5% level and 92-94% at the 10% level to reject the null hypothesis. This shows that our test is able to detect economically meaningful effects.

Table 4.5: Rejection Rates with Different Effect Sizes

		Rejection rate					
Size of change in belief difference	Equivalent to relative loss (if incentivized)	Main		Archive		No-Incentive	
		5% significance level	10% significance level	5% significance level	10% significance level	5% significance level	10% significance level
8	4%	0.50	0.64	0.53	0.67	0.52	0.66
10	6.25%	0.69	0.81	0.74	0.84	0.71	0.82
12	9%	0.85	0.92	0.88	0.94	0.86	0.93
14	12.25%	0.95	0.97	0.96	0.99	0.96	0.98

4.4 Conclusion and Discussions

In this paper, we examine the effect of making a choice on subsequent belief formation. Regardless of whether belief elicitation is incentivized or not, we find no evidence for a belief change, despite of the sufficient power of our test to detect a non-trivial effect. This suggests that making a choice does not lead to an economically meaningful belief change in our setting. We do not detect a stake-induced belief distortion, either. This suggests that wishful thinking does not drive an optimism bias in our setting.

Our results on state-induced belief distortion are consistent with some previous studies (Barron 2018; Heger and Papageorge 2018; the medium- and high-incentive group of Coutts 2018), but are different from findings of some other studies (Mayraz 2014; and the low-incentive group of Coutts 2018). We consider two possible reasons for the difference. First, the ratio of the stake in a good state to the incentive for belief elicitation might matter. The low-incentive group of Coutts (2018) has a ratio of 26.7, which is much higher than 8 in the medium-incentive group, 4 in the low-incentive group, 4.95 in Barron (2018) and one in Heger and Papageorge (2018). A higher ratio implies greater importance of anticipatory utility relative to accuracy, which induces people to make a larger distortion. Second, the salience of distortion may also be affected by whether the event about which beliefs are formed is immediately verifiable. Mayraz (2014) asks subjects to predict asset prices in one month. In that case, the event is not verifiable when subjects report their beliefs, and hence subjects could have less fear of negative evaluation from the experimenter (Watson and Friend 1969; see Trautmann et al. 2008 for an application in ambiguity aversion). This implies a smaller mental cost for subjects to distort their beliefs.

Our results can also be compared to the literature of cognitive dissonance. Most studies on cognitive dissonance provide evidence of choice-induced preference change, while our study does not find support for choice-induced belief change. One reason for the gap could be that preferences are always private information and thus cannot be directly verified. As discussed above, unverifiability makes preferences more likely to be distorted. The discrepancy can also arise from the difference between beliefs over attributes and preferences over alternatives. As mentioned in the introduction, preferences over alternatives can be divided into two parts: beliefs over attributes and preferences over lotteries of attributes. When more than one attribute is involved, it is possible that making a choice changes the relative importance of different attributes (e.g. health benefits become more important relative to deliciousness) but does not change beliefs about any attribute (e.g. beliefs about health benefits and deliciousness of all dessert options keep constant), and thereby the person displays a choice-induced preference change but not a belief change.

BIBLIOGRAPHY

Abdellaoui, M., Attema, A. E., & Bleichrodt, H. (2010). Intertemporal Tradeoffs for Gains and Losses: An Experimental Measurement of Discounted Utility. *Economic Journal, 120*(545), pp. 845-866.

Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2013a). Sign-Dependence in Intertemporal Choice. *Journal of Risk and Uncertainty, 47*(3), pp. 225-253.

Abdellaoui, M., Bleichrodt, H., l'Haridon, O., & Paraschiv, C. (2013b). Is There One Unifying Concept of Utility? An Experimental Comparison of Utility Under Risk and Utility Over Time. *Management Science, 59*(9), pp. 2153–2169.

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats etAxiomes de l'Ecole Americaine. *Econometrica, 21*(4), pp. 503-546.

Alós-Ferrer, C., & Shi, F. (2015). Choice-Induced Preference Change and the Free-Choice Paradigm: A Clarification. *Judgment and Decision Making, 10*(1), pp. 34-49.

Andersen, S., Cox, J. C., Harrison, G. W., Lau, M., Rutström, E. E., & Sadiraj, V. (2012). Asset Integration and Attitudes to Risk: Theory and Evidence. *Working Paper 2012-07*.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2011). Discounting Behavior and the Magnitude Effect. *Working Paper 2011-05*.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2013). Discounting Behavior and the Magnitude Effect: Evidence from a Field Experiment in Denmark. *Economica, 80*(320), pp. 670-697.

Andreoni, J., & Sprenger, C. (2012a). Estimating Time Preferences from Convex Budgets. *American Economic Review, 102*(7), pp. 3333-3356.

Andreoni, J., & Sprenger, C. (2012b). Risk Preferences Are Not Time Preferences. *American Economic Review, 102*(7), pp. 3357-3376.

Andreoni, J., Kuhn, M. A., & Sprenger, C. (2013). On Measuring Time Preferences. *NBER Working Paper No. 19392*.

Aronson, E., & Carlsmith, J. M. (1963). Effect of the Severity of Threat on the Devaluation of Forbidden Behavior. *Journal of Abnormal and Social Psychology, 66*(6), pp. 584-588.

Aronson, E., & Mills, J. (1959). The Effect of Severity of Initiation on Liking for a Group. *Journal of Abnormal and Social Psychology, 59*(2), pp. 177-181.

Attema, A. E., Bleichrodt, H., Gao, Y., Huang, Z., & Wakker, P. P. (2016). Measuring Discounting without Measuring Utility. *American Economic Review, 106*(6), pp. 1476-1494.

Augenblick, N., Niederle, M., & Sprenger, C. (2015). Working Over Time: Dynamic Inconsistency in Real Effort Tasks. *Quarterly Journal of Economics, 130*(3), 1067-1115.

Barron, K. (2018). Belief Updating: Does the 'Good-News, Bad-News' Asymmetry Extend to Purely Financial Domains? *WZB Discussion Paper*.

Baucells, M., & Heukamp, F. H. (2012). Probability and Time Trade-Off. *Management Science, 58*(4), pp. 831-842.

Bem, D. J. (1965). An Experimental Analysis of Self-Persuasion. *Journal of Experimental Social Psychology, 1*(3), pp. 199-218.

Bem, D. J. (1972). Self-Perception Theory. In L. Berkowitz, *Advances in Experimental Social Psychology* (Vol. 6, pp. 1-62). New York: Academic Press.

Benhabib, J., & Bisin, A. (2005). Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions. *Games and Economic Behavior, 52*(2), pp. 460-492.

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-Bias, Quasi-Hyperbolic Discounting, and Fixed Costs. *Games and Economic Behavior, 69*(2), pp. 205-223.

Bernheim, B. D., & Rangel, A. (2004). Addiction and Cue-Triggered Decision Processes. *American Economic Review, 94*(5), pp. 1558-1590.

Brehm, J. W. (1956). Postdecision Changes in the Desirability of Alternatives. *Journal of Abnormal and Social Psychology, 52*(3), pp. 384-389.

Brunnermeier, M. K., & Parker, J. A. (2005). Optimal Expectations. *American Economic Review, 95*(4), pp. 1092-1118.

Chakraborty, A., Calford, E. M., Fenig, G., & Halevy, Y. (2017). External and Internal Consistency of Choices Made in Convex Time Budgets. *Experimental Economics, 20*(3), pp. 687-706.

Chen, M. K., & Risen, J. L. (2010). How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm. *Journal of Personality and Social Psychology, 99*(4), pp. 573-594.

Cheung, S. L. (2015a). Eliciting Utility Curvature and Time Preference. *Working Paper 2015-01, School of Economics, The University of Sydney.*

Cheung, S. L. (2015b). Comment on “Risk Preferences Are Not Time Preferences”: On the Elicitation of Time Preference under Conditions of Risk. *American Economic Review, 105*(7), pp. 2242-2260.

Cohen, A. R. (1962). A Dissonance Analysis of the Boomerang Effect. *Journal of Personality, 30*(1), pp. 75-88.

Cooper, J. (2007). *Cognitive Dissonance: Fifty Years of a Classic Theory.* London: SAGE.

Coutts, A. (2015). Testing Models of Belief Bias: An Experiment. *MPRA Paper*.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2017). The Robustness and Pervasiveness of Sub-Additivity in Intertemporal Choice. *Working Paper 2017-07-18*.

Eil, D., & Rao, J. M. (2011). The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics, 3*(2), pp. 114-138.

Epper, T. (2015). Income Expectations, Limited Liquidity, and Anomalies in Intertemporal Choice. *University of St. Gallen, School of Economics and Political Science Discussion Paper No. 2015-19*.

Epstein, L. G., & Stanley, E. Z. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica, 57*(4), pp. 937-969.

Ertac, S. (2011). Does Self-Relevance Affect Information Processing? Experimental Evidence on the Response to Performance and Non-Performance Feedback. *Journal of Economic Behavior and Organization, 80*(3), pp. 532-545.

Festinger, L. (1957). *A Theory of Cognitive Dissonance.* Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive Consequences of Forced Compliance. *Journal of Abnormal and Social Psychology, 58*(2), pp. 203-210.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics, 10*(2), 171-178.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature, 40*(2), pp. 351-401.

Fudenberg, D., & Levine, D. K. (2006). A Dual-Self Model of Impulse Control. *American Economic Review, 96*(5), pp. 1449-1476.

Fudenberg, D., & Levine, D. K. (2011). Risk, Delay, and Convex Self-Control Costs. *American Economic Journal: Microeconomics, 3*(3), pp. 34-68.

Gerard, H. B., & Mathewson, G. C. (1966). The Effects of Severity of Initiation on Liking for a Group: A replication. *Journal of Experimental Social Psychology, 2*(3), pp. 278-287.

Halevy, Y. (2015). Time Consistency: Stationarity and Time Invariance. *Econometrica, 83*(1), pp. 335-352.

Heger, S. A., & Papageorge, N. W. (2018). We Should Totally Open a Restaurant: How Optimism and Overconfidence Affect Beliefs. *Working Paper*.

Holcomb, J. H., & Nelson, P. S. (1992). Another Experimental Look at Individual Time Preference. *Rationality and Society, 4*(2), 199-220.

Holden, S. T., & Quiggin, J. (2017). Bounded Awareness and Anomalies in Intertemporal Choice: Zooming in Google Earth as Both Metaphor and Model. *Journal of Risk and Uncertainty, 54*(1), pp. 15-35.

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review, 92*(5), pp. 1644-1655.

Kerschbamer, R. (2015). The Geometry of Distributional Preferences and a Non-Parametric Identification Approach: The Equality Equivalence Test. *European Economic Review, 76*(May), pp. 85-103.

Kirby, K. N. (1997). Bidding on the Future: Evidence against Normative Discounting of Delayed Rewards. *Journal of Experimental Psychology: General, 126*, pp. 54-70.

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin Addicts Have Higher Discount Rates for Delayed Rewards. *Journal of Experimental Psychology: General, 128*, pp. 78-87.

Laibson, D., Repetto, A., & Tobacman, J. (2007). Estimating Discount Functions with Consumption Choices over the Lifecycle.

Loewenstein, G. (1987). Anticipation and the Valuation of Delayed Consumption. *Economic Journal, 97*(387), pp. 666-684.

Loewenstein, G., & Prelec, D. (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. *The Quarterly Journal of Economics, 107*(2), pp. 573-597.

Mayraz, G. (2014). Priors and Desires: A Model of Optimism, Pessimism, and Cognitive Dissonance. *Working Paper*.

Meier, S., & Sprenger, C. (2010). Present-Biased Preferences and Credit Card Borrowing. *American Economic Journal: Applied Economics, 2*(1), pp. 193-210.

Miao, B., & Zhong, S. (2015). Comment on “Risk Preferences Are Not Time Preferences”: Separating Risk and Time Preference. *American Economic Review, 105*(7), pp. 2272-2286.

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2014). Managing Self-Confidence: Theory and Experimental Evidence. *Working Paper*.

Mullainathan, S., & Washington, E. (2009). Sticking with Your Vote: Cognitive Dissonance and Political Attitudes. *American Economic Journal: Applied Economics, 1*(1), pp. 86-111.

Noor, J. (2011). Intertemporal Choice and the Magnitude Effect. *Games and Economic Behavior, 72*(1), pp. 255-270.

Oja, H., & Randles, R. H. (2004). Multivariate Nonparametric Tests. *Statistical Science, 19*(4), pp. 598-605.

Read, D. (2001). Is Time-Discounting Hyperbolic or Subadditive? *Journal of Risk and Uncertainty, 23*(1), pp. 5-32.

Sun, C., & Potters, J. (2016). Magnitude Effect in Intertemporal Allocation Tasks. *Working Paper 2016-10, Department of Economics and CentER, Tilburg University*.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., & Trautmann, S. T. (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. *American Economic Review, 103*(1), pp. 510-531.

Thaler, R. (1981). Some Empirical Evidence on Dynamic Inconsistency. *Economics Letters, 8*(3), pp. 201-207.

Trautmann, S. T., Vieider, F. M., & Wakker, P. P. (2008). Causes of Ambiguity Aversion: Known versus Unknown Preferences. *Journal of Risk and Uncertainty, 36*(3), pp. 225-243.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science, New Series, 211*(4481), pp. 453-458.

Watson, D., & Friend, R. (1969). Measurement of Social-Evaluative Anxiety. *Journal of Consulting and Clinical Psychology, 33*(4), pp. 448-457.